

OM of: US-09-294-298-4 to: GenEmbl.* out_format : pfs

Date: Jan 19, 2001 12:44 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+2.p2n.model -DEV=slp
-Q=/cgn2_1/USPTO.spool/US09294298/runat_17012001_125501_27756/app_query.fasta_1.4359
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOPEXT=0.000 -QGAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-GAPEXT=7.000 -XGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEX=7.000 -START=1 -MATRIX=blossum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -OUTSCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=200000000 -USER=US09294298 -CGN1_13912 -NCPU=6
-ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-294-298-4

Query length: 1376

Database: GenEmbl.*

Database sequences: 1118133

Database length: -1736092196

Search time (sec): 10440.680000

score_list:

Sequence	Sfrd	Orig	ZScore	Escore	Len	Documentation
gb_ro:AF058790	+	7216.50	2903.19	1.5e-153	4140	AF058790 Rattus norvegicus Syn
gb_ro:AF058789	+	6980.50	2808.44	2.8e-148	4539	AF058789 Rattus norvegicus Syn
gb_ro:AF050183	+	6894.00	2774.21	2.3e-146	4265	AF050183 Rattus norvegicus GTF
gb_ro:AF048976	+	6415.00	2583.72	1.6e-135	4063	AF048976 Rattus norvegicus syn
gb_ro:AF048976	+	6308.50	2539.14	2.0e-133	4801	AB016962 Rattus norvegicus mRN
gb_hg20:AL161903	-	4066.00	1625.80	2.1e-82	150956	AL161903 Homo sapiens chromo
gb_pr7:HSJ570F3	-	4060.50	1626.05	2.1e-82	94770	AL050332 Homo sapiens sequence
gb_hg20:AL161906	+	3987.00	1593.63	1.3e-80	163923	AL161906 Homo sapiens chromo
gb_pr3:AF047711	+	2603.00	1060.08	6.9e-51	4368	AF047711 Homo sapiens ncgp mRN
gb_hg7:AC019800	-	1263.50	511.55	2.5e-20	56366	AC019800 Drosophila melanoga
gb_hg4:AC012162	+	1263.50	505.11	5.6e-20	191504	AC012162 Drosophila melanoga
gb_in1:AE003506	+	1263.50	505.11	7.6e-20	300994	AE003506 Drosophila melanoga
gb_hg4:AC012161	-	1257.00	502.41	7.9e-20	202741	AC012161 Drosophila melanoga
gb_in1:AB011280	-	1172.00	488.97	4.4e-19	3984	AB011280 Caenorhabditis elegans
gb_pr4:AK024488	+	1103.50	461.22	1.6e-17	4287	AK024488 Homo sapiens mRNA for
gb_pr4:AK024488	+	1088.50	458.32	3.0e-17	2385	AK024488 Homo sapiens cDNA FLJ
gb_hg6:AC016523	+	1082.00	438.23	3.0e-16	65961	AC016523 Drosophila melanoga
gb_ro:AF053938	+	1065.00	454.50	3.7e-17	829	AF053938 Rattus norvegicus p135
gb_hg22:AL365274	-	710.00	284.58	1.1e-07	172027	AL365274 Homo sapiens chromo
gb_in2:CEC0785	+	676.50	278.93	2.2e-07	39660	246266 Caenorhabditis elegans
gb_pr6:HS593C16	-	650.50	264.72	1.4e-06	81971	AC035702 Human DNA sequence
gb_hg3:AC011492	+	613.00	246.29	1.5e-05	157803	AC027051 Homo sapiens chromo
gb_hg17:AC073765	+	600.00	238.94	3.8e-05	237619	AC073765 Mus musculus clone
gb_in1:AB011129	+	502.00	228.29	0.0002	2276	AB011279 Caenorhabditis elegans
gb_hg3:AC011703	+	502.00	201.89	0.0043	159777	AC011703 Drosophila melanoga
gb_pr7:AB011110	+	488.50	214.18	0.0009	5579	AB011110 Homo sapiens mRNA for
gb_hg13:AC040900	+	479.50	193.05	0.0135	155332	AC040900 Homo sapiens chromo
gb_hg20:AL161906	-	479.00	192.39	0.0147	169323	AL161906 Homo sapiens chromo
gb_ro:RATGAP1	+	455.00	203.57	0.0035	3296	L13151 Rat GTPase-activating p
gb_pr8:HUNGAPB	+	437.50	196.33	0.0089	3456	M23612 Human GTPase-activating
gb_com:BTGAPR	+	437.50	195.17	0.0103	4307	M23379 Human GTPase-activating
gb_com:BTGAPR	+	433.50	193.37	0.0130	4473	X12602 Bovine mRNA for GTPase
gb_com:BTGAPR	+	433.00	195.77	0.0095	2730	U30857 Bos taurus R-Ras GTPase
gb_p12:BG272158	+	430.50	194.03	0.0119	3144	A272158 Blumeria graminis f.
gb_pat1:EL13125	+	429.50	194.54	0.0112	2647	EL13125 Bovine gene for R-Ras p
gb_in3:DM6012609	+	420.00	189.27	0.0219	3502	EJ132609 Drosophila melanoga
gb_hg12:AC027051	-	417.00	168.04	0.0341	156700	AC027051 Homo sapiens chromo
gb_pr4:AK026441	+	400.50	182.63	0.0514	2813	AK026441 Homo sapiens cDNA: FL
gb_in1:AB011283	+	396.50	187.93	0.0260	760	AB011283 Caenorhabditis elegans
gb_ro:MMU20238	+	387.00	175.25	0.1325	4107	U20238 Mus musculus GTPase-acu

gb_pr7:HSINSP4BP + 381.00 174.80 0.1404 2837 ! X89399 Homo sapiens mRNA fo
gb_in1:AB011285 + 378.50 180.26 0.0697 833 ! AB011285 Caenorhabditis eleg
gb_ro:AF086714 + 371.00 170.31 0.2497 3117 ! AF086714 Mus musculus rasGA
gb_pr3:AF086713 + 362.50 166.86 0.3885 3148 ! AF086713 Homo sapiens rasGA
seq_name: gb_ro:AF058790

seq_documentation_block: 4140 bp mRNA ROD 09-OCT-1998
LOCUS AF058790 Rattus norvegicus SynGAP-b mRNA, complete cds.
DEFINITION AF058790
ACCESSION AF058790
VERSION AF058790.1 GI:3722228
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 4140)
AUTHORS Kim,J.H., Liao,D., Lau,L.F. and Haganir,R.L.
TITLE SynGAP: a synaptic RasGAP that associates with the PSD-95/SAP90
protein family
JOURNAL Neuron 20 (4): 683-691 (1998)
MEDLINE 88240917

REFERENCE 2 (bases 1 to 4140)
AUTHORS Kim,J.H. and Haganir,R.L.
TITLE Direct Submission
JOURNAL Submitted (09-APR-1998) Neuroscience, Johns Hopkins Medical
Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
REFERENCE 3 (bases 1 to 4140)
AUTHORS Kim,J.H. and Haganir,R.L.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1998) Neuroscience, Johns Hopkins Medical
Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA
REMARK Sequence update by submitter
COMMENT On Oct 9, 1998 this sequence version replaced gi:3065890.
FEATURES
Location/Qualifiers
1. 4140
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/tissue_type="hippocampus"
1. 3750
/note="neuronal RasGAP; contains PH domain, C2 domain, GAP
domain in the N-terminal region, and a T/SXV motif at the
C-terminus"
/codon_start=1
/product="SynGAP-b"
/protein_id="AAC63511.1"
/db_xref="GI:3722229"

BASE COUNT 902 a 1328 c 1182 g 728 t
ORIGIN

alignment_scores:
 Quality: 7216.50 Length: 1377
 Ratio: 5.245 Gaps: 1
 Percent similarity: 99.927 Percent identity: 99.927

alignment_block:
 US-09-294-298-4 x AF058790 ..

Align seg 1/1 to: AF058790 from: 1 to: 4140

1 GlyProProArgProArgGlnGlnValProValGlnValProCysPr 17
 |||||
 11 GGGCCCCCAGCCCGCCGCTCAGGGGCTCCGGTTTCAGGTTCTTGCCTGCC 60
 |||||
 17 oLeuLeuProThrAlaSerLeuSerAlaAlaAlaAlaLeuProAlaAlaP 34
 |||||
 61 CTTCTTCCACCGCCAGCCTCTCCGCCGCCGCTGCTCTTCTCTGCTT 110
 |||||
 34 heArgGlyAsnThrThrTrpValAlaArgGlyGlyArgValSerProGly 50
 |||||
 111 TCCGGGGGAATACCACTTGGGTCGCTCAGGAGGAAGAGTGTCTCCGGGG 159
 |||||
 51 GlyLysGlnTyrSerMetGluAlaAlaProAlaAlaProPheArgProSe 67
 |||||
 160 GGGAAACAGTACAGCATGGAGCCGCCGCCGCTGCGCCCTTCCGGCCCTC 209
 |||||
 67 rGlnGlyPheLeuSerArgLeuLysSerSerIleLysArgThrLys 84
 |||||
 210 GCAAGGCTTCTGAGCCGAGGCTAAAAAGCTCCATCAACAGGTACAAAGT 259
 |||||
 84 erGlnProLysLeuAspArgThrSerSerPheArgGlnIleLeuProArg 100
 |||||
 260 CACAACCCAACTTGACCCGACACACGCTTTCACAGATCTCCCTCCGCTGC 309
 |||||
 101 PheArgSerAlaAspHisAspArgAlaArgLeuMetGlnSerPheLysG 117
 |||||
 310 TTCGGAAGTGCTGACCATGACCGGCCGCCGCTGATGACAGAGCTTCAAGGA 359
 |||||
 117 uSerHisSerHisGluSerLeuLeuSerProSerSerAlaAlaGluAla 134
 |||||
 360 GTCTCACTCCCATGAGTCTCTGCTGAGTCCAGCCAGAGTCTCTGAGGCC 409
 |||||
 134 euGlnLeuAsnLeuAspGluAspSerIleLeuLysProValHisSer 150
 |||||
 410 TGGAGCTCAACCTGGATGAAGACTCCATATCAAGCCAGTACACAGCTCC 459
 |||||
 151 IleLeuGlyGlnGluPheCysPheGluValThrThrSerSerGlyThrLy 167
 |||||
 460 ATCTTGGCCAGGAGTTCCTGCTTTGAGGTAAACAACATCGTCTGGGACAA 509
 |||||
 167 sCysPheAlaCysArgSerAlaAlaGluArgAspLysTrpIleGluAsn 184
 |||||
 510 ATGTTTTGGCTGCTGCGCTGTCAGCCGAAAGGGACAAATGGATGAGAATC 559
 |||||
 184 euGlnArgAlaValLysProAsnLysAspAsnSerArgArgValAspAsn 200
 |||||
 560 TACAGAGGGCTGTGAACCCCAACAGGACACAGCCGCCGGGTAGATAAC 609
 |||||
 201 valLeuLysLeuTrpIleIleGluAlaArgGluLeuProProLysLysAr 217
 |||||
 610 GTGCTGAAACTGATGGATCATGAAGCTCGAGAGTGGCCCCCAGAGAGCG 659
 |||||
 217 gTyrtYrCysGluLeuCysLeuAspMetLeuTyrAlaArgThrThrS 234
 |||||
 660 ATATTACTCGAGTATGCTTGACGACATGCTCTATGACGAGCACTT 709
 |||||
 234 erLysProArgSerAlaSerGlyAspThrValPheTrpGlyGluHisPhe 250
 |||||
 710 CCAAGCCCGCTCAGCCTCAGGAGACACTGCTTTTGGGGGAGCACTTC 759
 |||||
 251 GluPheAsnAsnLeuProAlaValArgAlaLeuArgLeuHisLeuTyrAr 267
 |||||

|||||
 760 GAGTTTAAACACCTGCTGCTGCTCCGGGCGCTCGGCTGCATCTGTACCG 809
 |||||
 267 gAspSerAspLysLysArgLysLysAspLysAlaGlyTyrValGlyLeuV 284
 |||||
 810 TGACTCGACAAAAAGCGAAGAGGACAGGACAGGCTACGTTGGCCTGG 859
 |||||
 284 alThrValProValAlaThrLeuAlaGlyArgHisPheThrGluGlnTrp 300
 |||||
 860 TGACTGTTCCAGTGGCCACCTGCTGGGCGCCACTTCACAGAGCAGTGG 909
 |||||
 301 TyrProValThrLeuProThrGlySerGlySerGlySerGlyMetGlySe 317
 |||||
 910 TACCCCGTGACCTGCCAACAGGAAGTGGGGGCTCTGGGGGTATGGGCTC 959
 |||||
 317 rGlyGlyGlyGlyGlySerGlyGlySerGlyLysGlyLysGlyLysGlyG 334
 |||||
 960 GGGGGGAGGAGGGGGGTCAGGGGCGGCTCAGGGGGCAAGGGAAGAG 1009
 |||||
 334 lYcysProAlaValArgLeuLysAlaArgTyrGlnThrMetSerIleLeu 350
 |||||
 1010 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1059
 |||||
 351 ProMetGluLeuTyrLysGluPheAlaGluTyrValThrAsnHisTyrAr 367
 |||||
 1060 CCAATGGAGCTATATAAGAGGTTGCAAGATATGTACCAACACCTACCG 1109
 |||||
 367 gMetLeuCysAlaValLeuGluProAlaLeuAsnValLysGlyLysGluG 384
 |||||
 1110 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1159
 |||||
 384 luValAlaSerAlaLeuValHisIleLeuGlnSerThrGlyLysAlaLys 400
 |||||
 1160 AGTTCGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1209
 |||||
 401 AspPheLeuSerAspMetAlaMetSerGluValAspArgPheMetGluAr 417
 |||||
 1210 GACTTCCTTTTCAGACATGGCCATGTGAGAGGTAGACCGGTTTCATGGAGCG 1259
 |||||
 417 gGluHisLeuIlePheArgGluAsnThrLeuAlaThrLysAlaIleGluG 434
 |||||
 1260 GGAACACTCATATTCGCGGAGAACACGCTCGCCACTAAAGCCATAGAG 1309
 |||||
 434 luTyrMetArgLeuIleGlyGlnLysTyrLeuLysAspAlaIleGlyGlu 450
 |||||
 1310 AGTATATGAGACTGATTGGCCAGAAATACCTCAAGGATGCCATTTGGGGAG 1359
 |||||
 451 PheIleArgAlaLeuTyrGluSerGluGluAsnCysGluValAspProIl 467
 |||||
 1360 TTCATCCGGGCTCTGTATGAATCTGAGGAGAACTGTGAAGTAGACCCCAT 1409
 |||||
 467 eLysCysThrAlaSerSerLeuAlaGluHisGlnAlaAsnLeuArgMetC 484
 |||||
 1410 CAAGTGACAGCTGCCAGTCTGGCAGAGCACCAGGCCAACCTCGCGGATGT 1459
 |||||
 484 yScysGluLeuAlaLeuCysLysValValAsnSerHisCysValPhePro 500
 |||||
 1460 GCTGTGAGTTGGCCCTGTGCAAGGTGTCAACTCCCATTCGCTGTTCGCG 1509
 |||||
 501 ArgGluLeuLysGluValPheAlaSerTrpArgLeuArgCysAlaGluAr 517
 |||||
 1510 AGGAGCTGAGAGGAGGTGTTGTCATCATGCGGCTGCGCTGTGCGAGAGCG 1559
 |||||
 517 gGlyArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuA 534
 |||||
 1560 GGGCCGGGAGGACATTCCTGACAGGCTGATCAGCGCTCGCTCTTCTTCCTCG 1609
 |||||
 534 rgPheLeuCysProAlaIleMetSerProSerLeuPheGlyLeuMetGln 550
 |||||
 1610 GCTTCTCTCTGCGCGGCTCATGTGCGCCAGTCTGTTTGGACTGATGACG 1659
 |||||
 551 GluTyrProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysVa 567
 |||||

1660 GAGTACCCAGATGAGCAGACCTCAGAACCCCTCACCCCTCATCGCCCAAGGT 1709
 567 lileGlnAsnLeuAlaAsnPheSerLysPheThrSerLysGluAspPheL 584
 1710 TATCCAGAACCCTGGCCAACTTTCCCAAGTTTACCTCAAGAGGAGCTTCC 1759
 584 euGlyPheMetAsnGluPheLeuGluLeuGluTTPGlySerMetGlnGln 600
 1760 TGGGCTTCATGAACAGAGTTTGGAGCTGGAATGGGTTTATGACAGCAA 1809
 601 PheLeuTyrlleGluSerAsnLeuAspThrLeuThrAsnSerSerSerPh 617
 1810 TTCTTGATGATATCATATCAACCTGGACACACTGACCAACAGCAGCTTT 1859
 617 eGluGlyTyrlleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuL 634
 1860 TGAGGGCTACATAGACTTGGGCGGAGCTCTCCACACTTCACGCCCTGC 1909
 634 euTrpGluValLeuProGlnLeuSerLysGluAlaLeuLeuLysLeuGly 650
 1910 TCTGGAGGTGTCGCCAGCTCAGCAAGGAAGCCCTCCTGAAGCTGGGC 1959
 651 ProLeuProArgLeuLeuSerAspIleSerThrAlaLeuArgAsnProAs 667
 1960 CCGCTGCCCGGCTCTCAGCGACATCAGCAGCCCTGAGGAACCCCTAA 2009
 667 nileGlnArgGlnProSerArgGlnSerGluArgAlaArgSerGlnProM 684
 2010 CATCAAGAGCAGCGAGCGCCAGAGCGAGCGCGCTCGGTCTCAGGCCA 2059
 684 etValLeuArgGlyProSerAlaGluMetGlnGlyTyrMetMetArgAsp 700
 2060 TGGTGTGCGGGCGGTGACGCGAAATGCGAGGGCTACATGATGCGGGAC 2109
 701 LeuAsnSerSerIleAspLeuGlnSerPheMetAlaArgGlyLeuAsnSe 717
 2110 CTCAACAGCTCCATCGACCTTCAGTCCCTTCATGGCTCGAGGCTCAACAG 2159
 717 rSerMetAspMetAlaArgLeuProSerProThrLysGluLysProProp 734
 2160 CTCTATGAGCATGGCTCGGCTCCCTCCCTCCCAACCAAGAGAAACCCCGC 2209
 734 roProProGlyGlyLysAspLeuPheTyrlleValSerArgProPro 750
 2210 CGCCCTCCCGGTGGGGTAAAGACCTGTTCTATGTGAGCGGCCACCA 2259
 751 LeuAlaArgSerSerProAlaTyrlleCysThrSerSerAspIleThrG1 767
 2260 CTGGCCCGGTCTCCCGACATCTGCACGAGCAGCTCGGACATCACAGA 2309
 767 uProGluGlnLysMetLeuSerValAsnLysSerValSerMetLeuAspL 784
 2310 GCCGAGCAGAGATGCTCAGTGTCAACAAGAGTGTGTCATGCTGGACC 2359
 784 euGlnGlyaspGlyProGlyGlyArgLeuAsnSerSerSerValSerAsn 800
 2360 TGCAGGGCGACGGGCTGGGGCGGCTTAAACAGCAGTAGTGTTCACAC 2409
 801 LeuAlaAlaValGlyAspLeuLeuHisSerSerGlnAlaSerLeuThrAl 817
 2410 CTGGAGCTGTGGGGACCTGTTCACCTCAAGCCAGGCTTCACCTGACAGC 2459
 817 aAlaLeuGlyLeuArgProAlaProAlaGlyArgLeuSerGlnGlySerG 834
 2460 AGCCTTGGGGTGGCGGCTGACCTGCGGGCGGCTCTCCCAAGGGAGTG 2509
 834 lySerSerIleThrAlaAlaGlyMetArgLeuSerGlnMetClyValThr 850
 2510 GCTCTTCATCAGCAGCGCGGCTGCGGCTCAGCCAGATGGGTGTCACT 2559
 851 ThrAspGlyValProAlaGlnGlnLeuArgIleProLeuSerPheGlnAs 867
 2560 ACGGATGGTTCGCCCGCCAGCAACTGGCGCATTCCTCTTCTTCCAGAA 2609

867 nProLeuPheHisMetAlaAlaAspGlyProGlyProAlaGlyHisG 884
 2610 CCCTCTCTTCATATGCTGCGCATGGACAGGCGCCCGCCAGCAGGCCATG 2659
 884 lyGlySerSerGlyHisGlyProProSerSerHisHisHisHisHis 900
 2660 GAGGAGCAGTGGCATGGTCCACCTTCTCCCATCACCACACCCACCC 2709
 901 HisHisHisHisArgGlyGlyGluProProGlyAspThrPheAlaProPh 917
 2710 CATCACCATCACCAGGGGAGAACCCCGAGGGGACACTTTTGGCCCGTT 2759
 917 eHisGlyTyrlleSerLysSerGluAspLeuSerThrGlyValProLysPro 934
 2760 CCATGGCTATAGCAAGAGCAGGACCTCTCTACAGGGGTCTTAAAGCCC 2809
 934 roAlaAlaSerIleLeuHisSerHisSerSerSerAspGluPheGlyPro 950
 2810 CTGGGCTCTCATCTTCACAGCCACAGCTACAGTACATGATGTTGGACC 2859
 951 SerGlyThrAspPheThrArgArgGlnLeuSerLeuGlnAspAsnLeuG1 967
 2860 TCTGGTACTGATTTTACCGCTCGGACGCTCTCCTTCAGGACACCTACA 2909
 967 nHisMetLeuSerProProGlnIleThrIleGlyProGlnArgProAlaP 984
 2910 GCACATGCTCTCCCGCCCGACATCACCATCGGTCCCGAGAGGCCAGCTC 2959
 984 roSerGlyProGlyGlySerGlyGlySerGlyGlyGlyGlyGlyGly 1000
 2960 CCTCAGGCCAGGAGGGGCGGTGGGGGCGAGTGGGGGGGGGGGGGG 3009
 1001 GlyGlnProProLeuGlnArgGlyLysSerGlnGlnLeuThrValSe 1017
 3010 GGCCAGCACCTCCCTTGCAGAGGGGCAAACTCTCAGCAGTTGACAGTGAG 3059
 1017 rAlaAlaGlnLysProArgProSerSerGlyAsnLeuLeuGlnSerProG 1034
 3060 TGCTGCCCAAGAACCCCGGCTCCAGCGGGAACCTATTGTCAGTCCCGGG 3109
 1034 luProSerTyrlleGlyProAlaArgProArgGlnGlnSerLeuSerLysGlu 1050
 3110 AACCAAGTTATGTTGCTGCGCTCCAGCGCAACAGAGCTTCAGCAAGAG 3159
 1051 GlySerIleGlySerGlySerGlySerGlyGlyGlyGlyGlyGly 1067
 3160 GGCAGCATTTGGGGCAGCGGGGCGACGGTGGCGAGGGGGTGGGGGCT 3209
 1067 uLysProSerIleThrLysGlnHisSerGlnThrProSerThrLeuAsnP 1084
 3210 CAAGCCCTCCATCACCAGCAGCATTTCCAGACATCCATCCAGCTGAACC 3259
 1084 roThrMetProAlaSerGluArgThrValAlaTrpValSerAsnMetPro 1100
 3260 CCACATGCGGCTCGGACGACCTGTAGCTGGGTGTCATATATGCT 3309
 1101 HisLeuSerAlaAspIleGluSerAlaHisIleGluArgGluGluTyrl 1117
 3310 CACCTGTCCGCTGACATCGAGAGTGACACATTTGAGCGGGAAGAGTACAA 3359
 1117 sLeuTyrlleGlySerLysSerMetAspGluSerArgLeuAspArgValL 1134
 3360 GCTGAAGAGTACTTCGAAGTCCATGAGGAGGAGCCGCTGGACAGGGTCA 3409
 1134 ysGlyTyrlleGluGluIleHisSerLeuLysGluArgLeuHisMetSer 1150
 3410 AGGAGTACGAGGAGAGATCCACTCTCAAGGAAAGGCTACACATGTCC 3459
 1151 AsnArgLysLeuGluTyrlleGluArgLeuLeuSerGlnGluGluL 1167
 3460 AACCGAAGCTGGAAGAGTACGAGCGGAGGCTGCTGTCTCCAGGAAGACA 3509

seq name: qb TO:AF058789

74 tgluLysSerSerIleLysArgThrLysSerGlnProLysLeuAspArg 90
|||||
362 GCTTAAAGCTCCATCAACAGTACAAAGTACACCCCAACTTGACCGG 411
91 thrSerSerPheArgGlnIleLeuProArgPheArgSerAlaAspHisAs 107
|||||
412 ACCAGCAGCTTTCGACAGATCCCTGCCTCGCTCCGAAAGTCTGACCATGA 461
107 pArgAlaArgLeuMetGlnSerPheLysGluSerHisSerHisGluSerL 124
|||||
462 CGGGCCCGCTGATGCAGAGCTTCAGGAGTCTCACTCCCATGAGTCCC 511
124 euLeuSerProSerSerAlaAlaGluAlaLeuGluLeuAsnLeuAspGlu 140
|||||
512 TGCTGAGTCCAGCAGTGCCTGCTGAGGCCCTGGAGCTCAACCTGGATGAA 561
141 AspSerIleLysProValHisSerSerIleLeuGlyGlnGluPheCy 157
|||||
562 GACTCCATATTCAAGCCAGTACACAGCTCCATCTGGGCCAGGAGTTCTG 611
157 spheGluValThrThrSerSerGlyThrLysCysPheAlaCysArgSerA 174
|||||
612 CTTTGAGGTACACATCGCTCTGGACAAAATGTTTGGCTGCTCGGCTG 661
174 laAlaGluArgAspLysTrpIleLysLeuGlnArgAlaValLysPro 190
|||||
662 CAGCCGAAAGGGAGCAAAATGGATTGAGAAATCTACAGAGGGCTGTGAAACCC 711
191 AsnLysAspAsnSerArgArgValAspAsnValLeuLysLeuTrpIleI 207
|||||
712 AACAGGACACACGCCCGGGGTAGATACGTGCTGAAACTATGGATCAT 761
207 eGluAlaArgGluLeuProProLysLysArgTyTrpCysGluLeuCysL 224
|||||
762 AGAAGCTCGAGAGTGCCTCCCGCCAGAGCGATATTACTCGGAGTTATGCC 811
224 euAspAspMetLeuTrpAlaArgThrThrSerLysProArgSerAlaSer 240
|||||
812 TSGAGACATGCTCTATGACGAGGACACCTTCCAAAGCCCGCTCAGCCTCA 861
241 GlyAspThrValPheTrpGlyGluHisPheGluPheAsnLeuProAl 257
|||||
862 GGAGACATGCTCTTTGGGGCGAGCACTTCGAGTTTAAACACTGCCTGC 911
257 aValArgAlaLeuArgLeuHisLeuTrpArgAspSerAspLysLysArgL 274
|||||
912 TGTCCGGCGCTGCGGCTGCATCTGTACCGTGACTCGGACAAAAGCGGA 961
274 yLysAspLysAlaGlyTyTrpValGlyLeuValThrValProValAlaThr 290
|||||
962 AGAAGGACAAAGCGAGGCTACGTTGGCTGTGTGACTGTTCCAGTGGCCACC 1011
291 LeuAlaGlyArgHisPheThrGluGlnTrpTrpProValThrLeuProTh 307
|||||
1012 CTGGCTGGCGCCACTTCACAGACAGTGTGTACCCGTGACCTGCAAC 1061
307 rGlySerGlySerGlyMetGlySerGlyGlyGlyGlyGlySerG 324
|||||
1062 AGAAGTGGGGCTCTGGGGTATGGGCTCGGGGGAGGAGGGGGGTGAG 1111
324 lycGlySerGlyGlyLysGlyLysGlyGlyCysProAlaValArgLeu 340
|||||
1112 GGGCGGCTCAGGGGGGAAAGGAGAGGCTGTCTGCTGCTGGCGGTG 1161
341 LysAlaArgTyTrpGlnThrMetSerIleLeuProMetGluLeuTyrlsG 357
|||||
1162 AAGGCCGTTTACCAGACATGAGTATCTCCCATGGAGCTATATAGA 1211
357 uPheAlaGluTrpValThrAsnHisTyTrpArgMetLeuCysAlaValLeuG 374
|||||
1212 GTTTCAGAAATATGTGACAAACCACTACCGCATGCTGTGTGCGGTGCTGG 1261

374 luProAlaLeuAsnValLysGlyLysGluValAlaSerAlaLeuVal 390
|||||
1262 AGCCCGCCCTCAATGTCAAGGGCAAGGAGGTCCCTAGTGCACGTGTT 1311
391 HisIleLeuGlnSerThrGlyLysAlaLysAspPheLeuSerAspMetAl 407
|||||
1312 CACATCTGCAAGACACAGCAAGGCCAAGGACTTCCCTTTCAGACATGGC 1361
407 aMetSerGluValAspArgPheMetGluArgGluHisLeuIlePheArgG 424
|||||
1362 CATGTCAGAGGTAGACCGGTTCATGAGCGGGAACACCTCATATTCCGCG 1411
424 luAsnThrLeuAlaThrLysAlaIleGluGluLysMetArgLeuIleGly 440
|||||
1412 AGAACAGCTCGCCACTAAAGCCATAGAGATATATGAGACTGATTGGC 1461
441 GlnLysTrpLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTrpG 457
|||||
1462 CAGAAATACCTCAAGGATGCCATTGGGAGTTTCATCCGGGCTCTGTATGA 1511
457 uSerGluGluAsnCysGluValAspProIleLysCysThrAlaSerSerL 474
|||||
1512 ATCTGAGGAGAACTGTGAGTACACCCCATCAAGTGCACAGCGTCCAGTC 1561
474 euAlaGluHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCys 490
|||||
1562 TGGCAGAGCACCAAGGCCAACCTCGGATGTGCTGTGAGTTGGCCCTGTGC 1611
491 LysValValAsnSerHisCysValPheProArgGluLeuLysGluValPh 507
|||||
1612 AAGTGTGTCAACTCCCATTTCCGTTTCCCGAGGAGCTGAAGGAGGTGTT 1661
507 eAlaSerTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaA 524
|||||
1662 TGCAATATGGCGGCTCGCTGTGCAGAGCGGGCGGAGGACATGTGCTG 1711
524 spArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIle 540
|||||
1712 ACAGGCTGATCAGGCGCTCCTCTCTCGCTTCTCTGCCCGCCATC 1761
541 MetSerProSerLeuPheGlyLeuMetGlnGluTrpProAspGluGlnTh 557
|||||
1762 ATGTCGCCAGTCTGTTTGGACTGTATGAGAGTACCCAGATGAGCAGAC 1811
557 rSerArgThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnP 574
|||||
1812 CTCACGAACCTCACCTCATGCCAAGGTTATCCAGAACCTGGCCAACT 1861
574 heSerLysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPhe 590
|||||
1862 TTTCCAGTTTACCTCAAGAGGAGACTTCTGGGCTTCATGAACGAGTTT 1911
591 LeuGluLeuGluTrpGlySerMetGlnGlnPheLeuTrpGluIleSerAs 607
|||||
1912 CTGGAGCTGGAGTGGGTTCTATGCAGCAATTTCTGTATGAGATATCCAA 1961
607 nLeuAspThrLeuThrAsnSerSerPheGlyGlyTyrlleAspLeuG 624
|||||
1962 CCTGGACACACTGACCAACAGCAGCAGTTCCTGAGGGCTACATAGACTTG 2011
624 lyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuProGln 640
|||||
2012 GCCCGAGCTCTCCACACTTCAGCGCTCTGCTGGGAGGTGCTGCCCCAG 2061
641 LeuSerLysGluAlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSe 657
|||||
2062 CTCACAGGAAGCCCTCTGAGGCTGGCGGCTGCCCCGGCTCCTCAG 2111
657 rAspIleSerThrAlaLeuArgAsnProAsnIleGlnArgGlnProSerA 674
|||||
2112 GCACATCAGCAGCCCTGAGAACCTTACATCCAAAGGCGAGCCGACCC 2161
674 rgGlnSerGluArgAlaArgSerGlnProMetValLeuArgGlyProSer 690

```

|||||
2162 GCCAGAGGAGCGGCTCGTCTCAGCCCATGGTGTGGCGGGCCGTCA 2211
691 AlaGluMetGlnGlyTyrMetMetArgAspLeuAsnSerSerIleAspLe 707
2212 GCCGAGATGCGAGGCTACATGATGCGGGAGCCCTCAACAGCTCCATCGACCT 2261
707 uGlnSerPheMetAlaArgGlyLeuAsnSerSerMetAspMetAlaArgL 724
2262 TCAGTCCCTTCAATGCTCGAGGCGCTCAACAGCTCTATGGACATGGCTGCC 2311
724 euProSerProThrLysGluLysProProProProProGlyGlyGly 740
2312 TCCCTTCCCAACCAAGAGAAACCCCGCGCCCTCCCGGTGGGGT 2361
741 LysAspLeuPheTyrValSerArgProProLeuAlaArgSerSerProAl 757
2362 AAAGACCTGTCTATGTGAGCGCGCCACCACTGGCCCGGTCTCCGCCAGC 2411
757 aTyrCysThrSerSerSerAspIleThrGluProGluGlnLysMetLeuS 774
2412 ATACTGCACGAGCAGCTCGACATCACAGAGCGGAGCAGAGATGCTGA 2461
774 ervAlAsnLysSerValSerMetLeuAspLeuGlnGlyAspGlyProGly 790
2462 GTGTCAACAAGAGTGTCTCATGCTGGACCTGCAGGGCGACGGGCTGGG 2511
791 GlyArgLeuAsnSerSerSerValSerAsnLeuAlaAlaValIcylAspLe 807
2512 GSCCGCCTTAAACAGCAGTGTGTTTCCAACTGGCAGCTGTTGGGAGCT 2561
807 uLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProA 824
2562 GTTGCACTCAAGCAGGCTTCACTGACAGCAGCTTGGGGTTCGGCCCTG 2611
824 laProAlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAlaAla 840
2612 CACCTGCGCGGCGCTCTCCCAAGGAGTGCTCTTCCATCACAGCAGCC 2661
841 GlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGl 857
2662 GGCATGCGCCTCAGCCAGATGGGTGCTACTACGATGGTGTCCCGCCCA 2711
857 nGlnLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaA 874
2712 GCAACTGGGCATCCCTCTTCTTCCAGAACCCCTCTTCCATATGGCTG 2761
874 laAspGlyProGlyProProAlaGlyHisGlyGlySerSerGlyHisGly 890
2762 CCGATGGACCAAGGCCCCCAGCAGGCCATGGAGGGAGCAGTGGCCATGCT 2811
891 ProProSerSerHisHisHisHisHisHisHisHisHisHisHisHis 907
2812 CCACCTTCTCCATCACACACACACACACACACATCACACACAGGGGG 2861
907 yGluProProGlyAspThrPheAlaProPheHisGlyTyrSerLysSerg 924
2862 AGAACCCCGAGGGACACTTGTGCGCGGTTCATGGCTATAGCAAGAGCG 2911
924 luAspLeuSerThrGlyValProLysProProAlaAlaSerIleLeuHis 940
2912 AGGACCTCTCTACAGGGGTCCCTAAGCCCGCTGCGGCCCTCCATCTCAC 2961
941 SerHisSerTyrSerAspGluPheGlyProSerGlyThrAspPheThrAr 957
2962 AGCCACAGCTACATGATGATGTGGACCCCTCTGGTACTGATTTTACCCG 3011
957 GArgGlnLeuSerLeuGlnAspAsnLeuGlnHisMetLeuSerProProG 974
3012 TCGGCAGCTCTCACTTCAGGACAACCTTACAGCACATGCTCTCCCGCCC 3061
974 lnlleThrIleGlyProGlnArgProAlaProSerGlyProGlyGlyGly 990
|||||

```

```

3062 AGATCACCATCGGTCCCCAGAGGCCAGCTCCCTCAGGGCCAGAGGGGCGC 3111
991 SerGlyGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1007
3112 ACTGTGGGGGAGTGTGGGGGGGTGGGGGCCAGCCACCTCCCTTCGA 3161
1007 nArgGlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLysProArgP 1024
3162 GAGGGGCAATCTCAGCAGTTGACGTAGTGTGCCAGAAACCCCGGC 3211
1024 roSerSerGlyAsnLeuLeuGlnSerProGluProSerTyrGlyProAla 1040
3212 CGTCCAGCGGAACCTATTGACGTCCCGGAGCAACAAATATGGTCTGCC 3261
1041 ArgProArgGlnGlnSerLeuSerLysGluGlySerIleGlycylSerGl 1057
3262 COTCCACGGCAACAGAGCCTCAGCAAGAGGGCAGCATTTGGGGGAGCGG 3311
1057 yGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyG 1074
3312 GGGCAGCGGTGGCGAGGGGTGGGGGCTCAAGCCCTCCATCACCAGC 3361
1074 lnhHisSerGlnThrProSerThrLeuAsnProThrMetProAlaSerGlu 1090
3362 AGCATTCACAGCTCCATCCAGCTGAACCCACGATCGCGCTCGGAG 3411
1091 ArgThrValAlaIleTrpValSerAsnMetProHisLeuSerAlaAspIleGl 1107
3412 CGGACTGTAGCTGGGTGTCAATATGCTTCACTGCTCGCTGACATCGA 3461
1107 uSerAlaHisIleGluArgGluGluTyrLysLeuLysGluTyrSerLysS 1124
3462 GAGTGCACACATTGACGGGAGAGTACAGCTGAAGAGTACTCGAAGT 3511
1124 erMetAspGluSerArgLeuAspArgValLysGlyTyrGluGluGluLle 1140
3512 CCATGGAGGAGCGGCTGACAGAGGTGAAGGAGTACGAGGAGGAGATC 3561
1141 HisSerLeuLysGluArgLeuHisMetSerAsnArgLysLeuGluGlu 1157
3562 CACTCCTGAAGGAAGGCTACACATGTCCACCGGAAGCTGGAAGAGTA 3611
1157 xGluArgArgLeuLeuSerGlnGluGluGlnThrSerLysIleLeuMetG 1174
3612 CGAGGGAGGCTGCTGCCAGGAGAGCAGACACCAAGATCCTGATGC 3661
1174 lntYrGlnAlaArgLeuGluGlnSerGluLysArgLeuArgGlnGlnGln 1190
3662 AGTACCAAGCCCGCTGGAGCAGAGCGAGAAGCGCTTGAGGAGCAGCAG 3711
1191 ValGluLysAspSerGlnIleLysSerIleIleGlyArgLeuMetLeuVa 1207
3712 GTGGAGAGAGGACTCCAGATCAAGAGCATCTTGGCAGGCTGATGCTGT 3761
1207 lGluGluLeuArgArgAspHisProAlaMetAlaGluProLeuProG 1224
3762 GGAGGAGAGCTGCGCGGGAGCACCCCGCATGGCTGAGCGCTGCCTG 3811
1224 luProLysLysArgLeuLeuAspAlaGln..... 1233
3812 AACCCCAAGAGGCTGCTCGAGCCTCAGCTCCTCATCAGGTAATTCTCC 3861
1233 ..... 1233
3862 TGGTTCCGCTTTTGACCGGGGCGAGACACTGGGGGAGGTGACTCCGGA 3911
1234 ..... ArgGlySerPheProProTyrValGlnGlnThrArgVal. 1246
3912 CCACGTGAGGAGAGGAGCTTCCCGCTTGGGTCCCAACAAACCCCGCTGT 3961
1247 ..ArgTrpProHisGlyThrAlaTrpProGlnProHisProPro 1262
3962 GACGTGGCCCCACCTTGGAGACGGCTTGGCCCCCCCCCAGCCCCCCCCC 4011

```

1263 HisProGlyCysArgSerGlnArgThrAlaSerSerGlyThrProGlnTh 1279
 4012 CACCCGGCTCCAGATCACAGAACGCGAGTCCGGAACACCGCAGAC 4061
 1279 rThrSerProSerIleThrAspLeuLeuProCysAlaProTyrProG 1296
 4062 CACTAGCCACCCAGACATCACAGCTCTCCCTGTGCACCTACCCCG 4111
 1296 lyProProSerValThrAspLeuLeuProSerAlaProAspProGlyThr 1312
 4112 CCCCACCCAGCGTCACAGACCTCTCCAGTGCACCCGACCCCTGGAACA 4161
 1313 SerProThrThrArgThrGlyArgHisGlnGlyThrAlaGlyLeuSerPr 1329
 4162 TCACCAACCCAGCAGTGGAGCTCACCAGGAGCAGCGGATGTCTCC 4211
 1329 oLeuThrProProThrGlyThrHisLeuSerThrProLeuLeuHisSerA 1346
 4212 CTACAGCGCTCTTGGGCGCAGCTCTGTCAACCCCACTGCTCCATCCA 4260
 1346 rArgGlySerGlyThrLeuSerCysProLeuThrProGlyHisHisLeu 1362
 4261 GGAGGAGAGTGGACCTCAGCTGCCCTCACCAGGAGCAGCCACCTA 4310
 1363 ProHisThrAspProPheThrLeuGlyCysTyrProHisPro 1376
 4311 CCCCACACAGCCCTTCACTCTGGGTGCTATCCCATCCT 4352

seq_name: gb_ro:AF050183

seq_documentation_block: 4265 bp mRNA ROD 15-SEP-2000
 LOCUS AF050183 Rattus norvegicus GTPase activating protein Syngap-c mRNA, complete
 DEFINITION cds.

ACCESSION AF050183

VERSION AF050183.2 GI:10140854

KEYWORDS Norway rat.

SOURCE Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .4265

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

/tissue_type="hippocampus"

/dev_stage="adult"

277. .3792

/note="neuronal synaptic RasGAP; associates with the

PSD-95/SAP90 protein family; alternatively spliced

variants, Syngap-a and Syngap-b"

/codon_start=1

/product="GTPase activating protein Syngap-c"

/protein_id="AAC40082.2"

/db_xref="GI:10140855"

/translation="MQSFKEHSHESLLSPSSAAALENLDEDSIIKPVHSSILGQE

FCFEVTSSTGTCFACRKAARERKQWLENLQRAVKPNKNSRDVNLKLTITAEALP

PKRYYCELCDDMLYATTSKPRASGDTVFWEHFEFNLPVAVRALRLHLRYSDK

KRKDKAGYVGLVTPVATLAGRHFTQWYPTVLTGSGSGSGSGSGSGSGSGSGSG

CDS

KKGGCPAVRLKARYQTMSILPMELYKEFAEYVTNHYRMLCAVLEPALNKGKEVAS
 ALVHILQSTGKAKDFLSDMAMSEVDREMERHILIFRENTLATKAIHEYNRLICOKYLK
 DAIGFIRALYSENECEVDPIKTASSLAEOHNLRCMCELALCKVYNHSCVFPREL
 KEVPASRLRCAERREDIADRLISASLFLFCPAIMSPSLFGLMOEYPDROTSTRL
 TLIAKIONLANFSKFTSKEDFLGFNNEFLLEWGSQOFLVELSNLDTLTNSSFEG
 YIDGRELSTLHALLWVLPOLSKALLKGLPLRLSLDISIALRNPINQOPRSQSE
 RASQPVWLRGFSAEQGYMRDUNSSIDLSFWMARGLNSMDMARLPSPTKEKPPPP
 PPGKDLFYVSRPLARSPAYCTSSDITEPEQMLSVNKSVMISIDLOGGPGRL
 NSSVSNLAAGVLDLLHSSQASLTAAALGRAPAGRLSQSGSSITAAAGRLSQMGVTT
 DGVPAQRLRIFLSENFLEHMAADGPGFAGHGGSGHGGPPSSHHHHHHHHHRRGEP
 PGDTFAPFHGYSKESDLSGVKPPAASILHSHSYDEFPGCTDFTRRQLSLQNLQ
 HMLSPQITIGPQRPAPSGGGGGGGGGGQPPPLQKRSOQLTVSAAQKPRPS
 SCNLIQSPSPGPAPRPOOSLSKESIGSGGGGGGGGGGLKPSITKOHSTPTLAN
 PTMPASERTVAVSNMPLHLSADTESAHIREEYKLEKYSKMDSEPLDRVKEVEIIEH
 SLKERLHMSNRKLEYERRLLSQEOTSKILMOYQARLEQSEKRLRQOQVEKDSQIKS
 IIGRLMVEEELRHPAMEPLPEPKAKRLDAQVERQLPPLPTNPRTVLAPPWNLG
 APPAPPPPPPIQITENGEFRNTADH"

BASE COUNT 936 a 1358 c 1203 g 768 t
 ORIGIN

alignment_scores:

Quality: 6894.00

Ratio: 5.168

Percent Similarity: 97.515

Percent Identity: 96.787

alignment_block:

US-09-294-298-4 x AF050183

Align seg 1/1 to: AF050183 from: 1 to: 4265

13 GlnValProCysProLeuLeuProThrAlaSerLeuSerAlaAlaAla 29
 28 CAGGCGCTTGTGCC...TTTCTCTCTGCCACGTTCTTCCAGGCTTC 74
 29 aLeuProAla...AlaPheArgGlyAsnThrThrTrpValAlaAaGlyG 45
 75 CCTGCCAGCTCAGGCTGTCTAGGAGAGACT..... 105
 45 lYArgValSerProGlyGlyGlnTyrSerMetGluAlaAlaProAla 61
 106GAGCTGCCCAACACCCCAT 126
 62 AlaProPheArgProSerGlnGlyPheLeuSerArgArgLeuLysSerSe 78
 127 TTCCCATTTTGTCTCCCGCAGAGCTTCTTGAGCGGAGGCTAAAAAGCTC 176
 78 rIleLysArgThrLysSerGlnProLysLeuAspArgThrSerSerPheA 95
 177 CATCAACGTACAAAGTCACACCCCAACTTGACCGGACCCAGCAGCTTC 226
 95 rGlnIleLeuProArgPheArgSerAlaAspHisAspArgAlaArgLeu 111
 227 GACAGATCTTCCCTCGCTTCCGAAGTGTGACCATGACCGGCGCGCTG 276
 112 MetGlnSerPheLysGluSerHisSerHisGluSerLeuLeuSerProSe 128
 277 ATGCAGAGCTTCAGGAGTCTCCTCCCATGAGTCTCTGAGTCCAGCCAG 326
 128 rSerAlaAlaGluAlaLeuGluLeuAsnLeuAspGluAspSerIleIleL 145
 327 CAGTGTCTGCTGAGGCGCTGAGGCTCAACCTGGATGAAGACTCCATTATCA 376
 145 ySProValHisSerSerIleLeuGlyGlnGluPheCysPheGluValThr 161
 377 AGCCAGTACAGCTCCTCCTGGCAGAGTCTGCTTGGTGAAGTGAACA 426
 162 ThrSerSerGlyThrLysCysPheAlaCysArgSerAlaAlaGluArgAs 178
 427 ACATCGTCTGGCAAAATGTTTTCCTGCTGCTGCTGCGGCGGAAAGGA 476
 178 pLysTrpIleGluLeuGlnArgAlaValLysProAsnLysAspAsnS 195

477 CAATGATTGAGATCTACAGAGGGCTGTGAACCCCAAGGACACA 526
 195 erArgArgValAspAsnValLeuLysLeuTrpIleIleGluAlaArgGlu 211
 527 GCCCGGGGTAGATAACGTGCTGAACATATGATCATAGAAGCTCGAGAG 576
 212 LeuProLysLysArgTyrTyrCysGluLeuLysLeuAspAspMetLe 228
 577 CTGCCCCCAAGAGCGATTACTTGCAGATTATGCTTGGACGACATGCT 626
 228 uTyrAlaArgThrThrSerLysProArgSerAlaSerGlyAspThrValP 245
 627 CTATGACGAGACACATTCCAAAGCCCGCTCAGCCTCAGAGACATGCT 676
 245 heTrpGlyGluHisPheGluPheAsnAsnLeuProAlaValAlaLeu 261
 677 TTTGGGGCGAGCACTTCGAGTTTAAACACCTGCTGCTCCGGGCGCTG 726
 262 ArgLeuHisLeuTyrArgAspSerAspLysLysArgLysAspLysAl 278
 727 CGGCTGCATCTGTACCTGACTCGGACAAAAGCGGAAGAGGACAAAGC 776
 278 agLyrTyrValGlyLeuValThrValProValAlaThrLeuAlaGlyArgH 295
 777 AGCTACGTTGGCTGGTGAAGTTTCCAGTGGCCACCTGCTGGTGGGCGCC 826
 295 isPheThrGluGlnTrpTyrProValThrLeuProThrGlySerGlyGly 311
 827 ACTTCACAGAGCAGTGGTACCCTGTCCTGCCAACAGGAAGTGGGGGC 876
 312 SerGlyGlyMetClySerGlyGlyGlyGlySerGlyGlySerGly 328
 877 TCTGGGGGTATGGCTCGGGGGAGAGGGGGTTCAGGGGGCGGCTCAGG 926
 328 YGlyLysGlyLysGlyCysProAlaValArgLeuLysAlaArgTyrG 345
 927 GGGCAAGGGAAGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
 345 InThrMetSerIleLeuProMetGluLeuTyrLysGluPheAlaGluTyr 361
 977 AGCAATGAGTATCTGCCCCTGAGGCTATATAGGAGTTTGCAAGATAT 1026
 362 ValThrAsnHisTyrArgMetLeuLysAlaValLeuGluProAlaLeuAs 378
 1027 GTACCAACCACTTACCGCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 1076
 378 nValLysGlyLysGluGluValAlaSerAlaLeuValHisIleLeuGln 395
 1077 TGTCAAGGCGAGGAGGAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1126
 395 erThrGlyLysAlaLysAspPheLeuSerAspMetAlaMetSerGluVal 411
 1127 GCACAGGCAAGGCAAGGACTTCTTTCAGACATGGCCATGCGGAGGTA 1176
 412 AspArgPheMetClyArgGluHisIleLeuPheArgGluAsnThrLeuAl 428
 1177 GACCGGTTCATGAGGGGGAACCTCATATTCGCGGAGAACGCTCGC 1226
 428 aThrLysAlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuL 445
 1227 CACTAAAGCCATAGAAGAGTATATGAGACTGATTGGCCAGAAATACCTCA 1276
 445 ysAspAlaIleGlyGluPheIleArgAlaLeuTyrGluSerGluGluAsn 461
 1277 AGGATGCCATTTGGGAGTTTCATCCGGGCTCTGTATGAATCTTGAGGAGA 1326
 462 CysGluValAspProIleLysCysThrAlaSerSerLeuAlaGluHisGly 478
 1327 TGTGAAGTAGACCCCATCAAGTGCACACGGCTCCAGTCTGGCAGAGACCA 1376
 478 nAlaAsnLeuArgMetCysGlyGluLeuAlaLeuCysLysValValAsnS 495
 1377 GGCACACCTGCGGATGTGCTGTGAGTTGGGCTGTGCAAGGTGTCAACT 1426

495 erHisCysValPheProArgGluLeuLysGluValPheAlaSerTrpArg 511
 1427 CCCATTGGCTGTTCCCGAGGAGGTGAAGAGGTGTTTGCATCATGGCGG 1476
 512 LeuArgCysAlaGluArgGlyArgGluAspIleAlaAspArgLeuIleSe 528
 1477 CTGGCTGTGCAGAGCGGGCCGGAGGACATTGCTGACAGGCTGATCAG 1526
 528 rAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetSerProSerL 545
 1527 CGCCTGCTCTTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1576
 545 euPheGlyLeuMetGlnGluTyrProAspGluGlnThrSerArgThrLeu 561
 1577 TGTGTGGACTGATGACGAGGTATCCAGAACCTGGCCAACTTTTCCAAGTTTAC 1626
 562 ThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerLysPheTh 578
 1627 ACCCTCATCGCAAGGTATCCAGAACCTGGCCAACTTTTCCAAGTTTAC 1676
 578 rSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGluLeuLut 595
 1677 CTCAAGAGGAGGACTCTCTGGCTTCATGAACGAGTTTCTGGAGCTGGAAT 1726
 595 rpGlySerMetGlnGlnPheLeuTyrGluIleSerAsnLeuAspThrLeu 611
 1727 GGGGTTCTATGCAGCAATTTCTGTATGAGATATCCAACTGGACACACTG 1776
 612 ThrAsnSerSerPheGluGlyTyrIleAspLeuGlyArgGluLeuSe 628
 1777 ACCAACAGCAGCAGTTTGGAGGCTACATAGACTTGGCCGCGAGCTCTC 1826
 628 rThrLeuHisAlaLeuLeuTrpGluValLeuProGlnLeuSerLysGluA 645
 1827 CACACTTCACGCCCTGCTCTGGAGGTGCTGCTGCTGCTGCTGCTGCTGCT 1876
 645 laLeuLysLysLeuGlyProLeuProArgLeuLeuSerAspIleSerThr 661
 1877 CCCCTCGAAGCTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1926
 662 AlaLeuArgAsnProAsnIleGlnArgGlnProSerArgGlnSerGluAr 678
 1927 GCCCTGAGGAACCTTAACATCAAGGACGCGGAGCGGCGGAGCGGCGG 1976
 678 gAlaArgSerGlnProMetValLeuArgGlyProSerAlaGluMetGlnG 695
 1977 CGCTCGGCTCAGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2026
 695 lyTyrMetMetArgAspLeuAsnSerSerIleAspLeuGlnSerPheMet 711
 2027 GCTACATGATGCGGACCTCAACAGCTCCATCGACCTTCAGTCTTCATG 2076
 712 AlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSerProTh 728
 2077 GCTCGAGGCTCAACAGCTCTATGGACATGGCTGCTGCTGCTGCTGCTGCT 2126
 728 rLysGluLysProProProProProGlyGlyGlyLysAspLeuPheT 745
 2127 CAAGGAAACCCCGCCCGCCCTCCCGTGGGGTAAAGACCTGTTCT 2176
 745 yrValSerArgProProLeuAlaArgSerSerProAlaTyrCysThrSer 761
 2177 ATGTGAGCCGCGCCACCTGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2226
 762 SerSerAspIleThrGluProGluGlnLysMetLeuSerValAsnLysSe 778
 2227 AGCTCGGACATCACAGAGCGGAGAGAGATGCTGAGTGTCAACAAGAG 2276
 778 rValSerMetLeuAspLeuGlnGlyAspGlyProGlyGlyArgLeuAsnS 795
 2277 TGTGTCCATGCTGAGACCTGAGGCGGACGGGCTGGGGCGGCGCTTAACA 2326

795 erSerSerValSerAsnLeuAlaValAlaGlyAspLeuLeuHisSerSer 811
 2327 GCAGTAGTGTTCACCACTGGGAGCTGTGGGGACCTGTGGCACTCAAGC 2376
 812 GlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaProAlaGlyAr 828
 2377 CAGGCTTCACTGACAGAGCCTTGGGGTGGGGCTGCACCTGCGGGGCG 2426
 828 gLeuSerGlnGlySerGlySerSerIleThrAlaAlaGlyMetArgLeuS 845
 2427 CCTCTCCCAAGGAGTGGCTCTTCCATCAGACAGCGCGGCATCGCTCA 2476
 845 erGlnMetGlyValThrThrAspGlyValProAlaGlnGlnLeuArgIle 861
 2477 GCCAGATGGGTGTCACACAGGATGGTGTCCCGCCAGCAACTGCGCATC 2526
 862 ProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaAspGlyProG 878
 2527 CCTCTTCTTCCAGAACCTCTCTTCCATATGGCTGCCATGGAGCAGG 2576
 878 yProProAlaGlyHisGlySerSerGlyHisGlyProProSerSerH 895
 2577 GCCCCACAGCGCATGGAGGAGCAGTGGCCATGGTCCACCTTCCTCCC 2626
 895 IsHisHisHisHisHisHisHisHisHisHisArgGlyGlyGluProGly 911
 2627 ATCACCACACACACACACATCACCAATCACCGAGGGGAGAACCCCGAGG 2676
 912 AspThrPheAlaProPheHisGlyTyrSerLysSerGluAspLeuSerTh 928
 2677 GACACTTTTGGCCGCTTCCATGGCTATAGCAAGAGCGAGGACCTCTCTAC 2726
 928 rGlyValProLysProProAlaAlaSerIleLeuHisSerHisTyrS 945
 2727 AGGGGTCTCTAAGCCCTCTGGGCTCTCATCTCTCACGCCACAGCTACA 2776
 945 erAspGluPheGlyProSerGlyThrAspPheThrArgArgGlnLeuSer 961
 2777 GTGATGAGTTTGGACCTCTGTGACTGATTTTACCCCTCGGCAGCTCTCA 2826
 962 LeuGlnAspAsnLeuGlnHisMetLeuSerProGlnIleThrIleG 978
 2827 CTTACAGGACACCTTACAGCATGCTCTCCCGCCCGCCAGATCACCATCG 2876
 978 yProGlnArgProAlaProSerGlyGlySerGlyGlyGlyGlyS 995
 2877 TCCCGAGGCGGAGCTCCCTCAGGGCCAGAGGGGAGTGGTGGGGCA 2926
 995 erGlyGlyGlyGlyGlyGlyGlnProProProLeuGlnArgGlyLysSer 1011
 2927 GTGGTGGGGGGGTGGGGGCCAGCCACCTCCCTTGCAGAGGGGCAATCT 2976
 1012 GlnGlnLeuThrValSerAlaAlaGlnLysProArgProSerSerGlyAs 1028
 2977 CAGCAGTGTACAGTAGTGTGCTCCAGAAACCCCGCCGCTCCAGCGGAA 3026
 1028 nLeuLeuGlnSerProGluProSerTyrGlyProAlaArgProArgGlnG 1045
 3027 CCTATTGACGTCCCGGAACCAAGTTATGGTCTCCCGCTCCAGCGCAAC 3076
 1045 InSerLeuSerLysGluGlySerIleGlyGlySerGlyGlySerGlyGly 1061
 3077 AGAGCCTCAGCAAGAGGCGCAGCATTTGGGGGCGAGCGGGGCGGTGGC 3126
 1062 GlyGlyGlyGlyGlyLeuLysProSerIleThrLysGlnHisSerGlnTh 1078
 3127 GGAGGGGTGGGGGCTCAAGCCTCCATCCCAAGCAGCATTCACAGAC 3176
 1078 rProSerThrLeuAsnProThrMetProAlaSerGluArgThrValAla 1095
 3177 TCCATCCACGTGAAACCCACGATGCGGCTCGGAGCGGAGCTAGCCT 3226
 1095 rpValSerAsnMetProHisLeuSerAlaAspIleGluSerAlaHisIle 1111

3227 GGGTGTCCATAATAGCTCACCCTGTCGCTGACATCGAGAGTGCACACTT 3276
 1112 GluArgGluGluTyrLysLeuLysGluTyrSerLysSerMetAspGluSe 1128
 3277 GAGCGGGAAGAGTACAACTGAAGAGTACTCGAAGTCCATGGACGAGAG 3326
 1128 rArgLeuAspArgValLysGluTyrGluGluIleHisSerLeuLysG 1145
 3327 CCGACTGGACAGGCTGAAGAGTACGAGGAGAGATCCACTCACTGAAGG 3376
 1145 luArgLeuHisMetSerAsnArgLysLeuGluTyrGluArgGluLeu 1161
 3377 AARGCTACACATGTCCAAACCGAAGCTGGAGAGAGTACGAGCGGAGGCTG 3426
 1162 LeuSerGlnGluGlnThrSerLysLeuLeuMetGlnTyrGlnAlaAr 1178
 3427 CTGTCCCGAGGAGAGCAGCAGCAAGATCTGTGATGAGTACCAAGCCCG 3476
 1178 gLeuGluGlnSerGluLysArgLeuArgGlnGlnValGluLysAspS 1195
 3477 CCTGGACAGCGCAGAGCGCTTGAGGCGCAGCAGGTGGAGAGGACT 3526
 1195 erGlnIleLysSerIleIleGlyArgLeuMetLeuValGluGluLeu 1211
 3527 CCCAGATCAAGAGCATCTTGGCAGCTGTGCTGGAGGAGGAGCTG 3576
 1212 ArgArgAspHisProAlaMetAlaGluProLeuProGluProLysLysAr 1228
 3577 CGCGGGACCCCGCCATGCTGAGCGCTGCTGAACCCCAAGAGAG 3626
 1228 gLeuLeuAspAla. Gln... ArgGlySerPheProProTyrValGlnGln 1243
 3627 GCTGCTGACGCTCAGTGGAGAGGAGCTTCCCGCTGGGTCCACAA 3676
 1244 ThrArgVal... ArgTrpProHisLeuGlyThrAlaTrpProGlnPr 1259
 3677 ACCGCGGTGTGAGCTGGCGCCACCTTGAACGGCTGGCGCCCGCCAGCC 3726
 1259 oHisProProHisProGlyCysArgSerGlnArgThrAlaSerSerGlyT 1276
 3727 CCACCCCGCCCGCGCTGCGAGATCAAGAGAACGGCGAGTTCCGGAA 3776
 1276 hrProGlnThrThrSerProProSerIleThrAspLeuLeuProCysAla 1292
 3777 CACGCGAGACCTAGCCACCCAGCATCACAGACCTCTTCCCTGTGCA 3826
 1293 ProTyrProGlyProProSerValThrAspLeuLeuProSerAlaProAs 1309
 3827 CCTTACCCCGCCCGCCAGCTCAGACCTCTTCCAGTGCACCCGA 3876
 1309 pProGlyThrSerProThrThrArgThrGlyArgHisGlnGlyThrAlaG 1326
 3877 CCTTGAACATCAACCAACCCAGGAGTGGAGCTCACCAGGGAGACGCG 3926
 1326 yLeuSerProLeuThrProProTyrGlyThrHisLeuSerThrProLeu 1342
 3927 GATTGTCTCTTAACTCTTGGGGCCCATCTGTCACCCCACTC 3976
 1343 LeuHisSerArgArgGluSerGlyThrLeuSerCysProLeuThrProG 1359
 3977 CTCTATTCCAGGAGAGTGGGAGCTTCCAGCTGCTTCCACCCCGAG 4026
 1359 yHisHisLeuProHisThrAspProPheThrLeuGlyCystYrProHisP 1376
 4027 ACACCACTACCCACACAGACCCCTTCACTCTGGGGTGTATCCCATC 4076
 1376 ro 1376
 4077 CT 4078

983 GGAGACACTGCTCTTTGGGGCGGACACTTCGAGTTTAAACAACCTGCCTGC 1032
257 aValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysArgL 274
1033 TCTCCGGGGCTGCGGGCTGCTGTATCCCTGACTCGGACAAAAGCGGA 1082
274 yLysAspLysAlaGlyTyrValGlyLeuValThrValProValAlaThr 290
1083 AGAAGGACACGACAGGCTACGTTGGCTGGTGACTGTTCCAGTGGCCACC 1132
291 LeuAlaGlyArgHisPheThrGluInTrpTyrProValThrLeuProTh 307
1133 CTGGCTGGGGCCACTTCACAGACAGTGTACCCCGTGACCTGCCAAC 1182
307 rGlySerGlySerGlyMetGlySerGlyGlyGlyGlyGlyGlyG 324
1183 AGGAAGTGGGGCTCTGGGGGTATGGGCTCGGGGGGAGGAGGGGGTCCAG 1232
324 yGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyG 340
1233 GGGGGCTCAGGGGGCAAGGGAAGAGGCTGCTCTGCTGGCGCTG 1282
341 LysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeuTyrLysG 357
1283 AAGGCCGCTTACCAACAGATGAGTATCTCCCATGGAGCTATATAAGGA 1332
357 uPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaValLeuG 374
1333 GTTTCAGAAATGTGACCAACCACTACCGCATGCTGTGTCGCTGTG 1382
374 luProAlaLeuAsnValLysGlyGlyGluGluValAlaSerAlaLeuVal 390
1383 AGCCGGCCCTCAATGTCGAAGGGCAAGGAGGAGGTGCTGTCACATG 1432
391 HisIleLeuGlnSerThrGlyLysAlaLysAspPheLeuSerAspMetAl 407
1433 CACATCTGCAAAAGCACAGCAAGGCCAAGGACTTCTCTTCAGACATGGC 1482
407 aMetSerGluValAspArgPheMetGluArgGluHisLeuIlePheArgG 424
1483 CATGTCAGAGTAGACCGGTTTCATGGAGCGGGAACACCTCATATTCG 1532
424 luAsnThrLeuAlaThrLysAlaIleGluGluTyrMetArgLeuIleGly 440
1533 AGAACACGCTCGCCACTAAAGCCATAGAAGAGTATATGAGACTGATGGC 1582
441 GlnLysTyrLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTyrG 457
1583 CAGAAATACCTCAAGGATGCCATTGGGGAGTTTCATCCGGGCTCTGTATGA 1632
457 uSerGluGluAsnCysGluValAspProIleLysCysThrAlaSerSerL 474
1633 ATCTGAGGAGAACTCTGAAGTAGACCCCATCAAGTGCACAGCTCCAGTC 1682
474 euAlaGluHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCys 490
1683 TGCGAGAGCACACAGGCCAACCTCGCGATGTGCTGTGAGTTGGCCCTGTGC 1732
491 LysValValAsnSerHisCysValPheProArgGluLeuLysGluValph 507
1733 AAGGTGGTCACTCCCATTCGCTGTCGCGAGGAGCTGAAGAGGTGTT 1782
507 eAlaSerTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaA 524
1783 TGCATCATGGCGCTGCTGTGCAGAGCGGGGGCGGGAGGACATTTGCTG 1832
524 spArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIle 540
1833 ACAGGCTGATACAGCGCTCGCTCTCTCCCTGCGCTTCTCTCCCGGCCATC 1882
541 MetSerProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnTh 557
1883 ATGTCGCCAGTCTGTTGGACTGATGACGAGGATACCCAGATGACGAGAC 1932

557 rSerArgThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnP 574
1933 CTCAGAACCTCACCTCATCCCAAGGTTATCCAGAACCTGGCCAAC 1982
574 heSerLysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPhe 590
1983 TTTCCAACTTTACCTCAAGGAGGAGACTTCTGGGCTTCATGAACGAGTTT 2032
591 LeuGluLeuGluTrpGlySerMetGlnGlnPheLeuTyrGluIleSerAs 607
2033 CTGGAGCTGGAGTGGGTTCTATGACGCAATTCCTGTATGACATATCAA 2082
607 nLeuAspThrLeuThrAsnSerSerSerPheGluGlyTyrIleAspLeuG 624
2083 CTTGACACACTGACCAACAGCAGCAGTTTGGAGGCTACATAGACTTGG 2132
624 yArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuProGln 640
2133 GCGCGAGCTCTCCACACTTCACGCCCTGCTGTGGAGGTGCTGCCCCAG 2182
641 LeuSerLysGluAlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSe 657
2183 CTCAGAGGAAGCCCTCTGAGCTGGGCCCTGCCCGCTCTCTCAG 2232
657 rAspIleSerThrAlaLeuArgAsnProAsnIleGlnArgGlnProSerA 674
2233 CGACATCAGCAGCAGCCCTGAGAAACCTTAACATCCAAAGCAGCGAGCC 2282
674 rGlnSerGluArgAlaArgSerGlnProMetValLeuArgGlyProSer 690
2283 GCCAGAGCAGCGGCTGCTGCTCAGCCCATGCTGCTGGCGGGCGCTCA 2332
691 AlaGluMetGlnGlyTyrMetMetArgAspLeuAsnSerSerIleAspLe 707
2333 GCGGAGATGCGAGGCTACATGATCGGGGACCTCAACAGCTCCATGACCT 2382
707 uGlnSerPheMetAlaArgGlyLeuAsnSerSerMetAspMetAlaArgL 724
2383 TCAGCTCTTCATGCTCGAGGCTTCAACAGCTCTATGGACATGGCTCGCC 2432
724 euProSerProThrLysGluProProProProProProGlyGlyGly 740
2433 TCCCTCCCAACCAAGGAGAAACCCCGCCGCTCCCTCCGCTGGGGGT 2482
741 LysAspLeuPheTyrValSerArgProProLeuAlaArgSerSerProAl 757
2483 AAAGACCTGTTCTATGTGAGCGGCCACCTGCGCGGCTCTCCCCAGC 2532
757 aTyrCysThrSerSerSerAspIleThrGluProGluGlnLysMetLeuS 774
2533 ATACTGCACGACAGCTCGGACATCACAGAGCGGAGCAGAGATGCTGA 2582
774 erValAsnLysSerValSerMetLeuAspLeuGlnGlyAspGlyProGly 790
2583 GTGTCAACAAGAGTGTCTCATGTGACCTGCGAGGCGGCGGCGCTGGG 2632
791 GlyArgLeuAsnSerSerSerValSerAspLeuAlaValGlyAspLe 807
2633 GCGCGCTTACACGAGTAGTGTTCACCTGGCAGCTGTGGGAGCT 2682
807 uLeuHisSerSerGlnAlaSerLeuThrAlaLeuLeuGlyLeuArgProA 824
2683 GTTGCACTCAAGCCAGGCTTCACTGACAGCAGGCTTGGGTTGCGGCTG 2732
824 laProAlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAlaAla 840
2733 CACTGCCGGGCGCTCTCCCAAGGAGTGGCTTCCATCAGCAGCAGCC 2782
841 GlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGl 857
2783 GGCATGCGCCTCAGCCAGATGGGTGTCTACACGAGTGGTGTCTCCGCCCA 2832

```

857 nGlnLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaA 874
|||||
2833 GCAACTGCGCATCCCTCTTTCTCCAGAACCCCTCTCTCCATATGCTG 2882
|||||
874 laAspGlyProGlyProProAlaGlyHisGlySerGlyHisGly 890
|||||
2883 CCGATGGACCGAGGGCCCGCCAGAGGCCATGGAGGAGCAGTGGCCATGCT 2932
|||||
891 ProProSerSerHisHisHisHisHisHisHisHisHisHisHis 907
|||||
2933 CCACCTTCTCCCATCACCAACCACCACCATCACCATCACCGAGGGG 2982
|||||
907 yGluProGlyAspThrPheAlaProPheHisGlyTySerLysSerC 924
|||||
2983 AGAACCCCGAGGGGACACTTTGCCCCGTTCCATGGCTATAGCAAGCG 3032
|||||
924 luAspLeuSerThrGlyValProLysProProAlaAlaSerIleLeuHis 940
|||||
3033 AGGACCTCTACAGGGGTCCCTAAGCCCTCGCGCTCCATCCCTCAC 3082
|||||
941 SerHisSerTySerAspGluPheGlyProSerGlyThrAspPheThrAr 957
|||||
3083 ASCACACGCTACGATGATGAGTTGGACCTCTGGTACTGATTTTACCG 3132
|||||
957 gArgGlnLeuSerLeuGlnAspAsnLeuGlnHisMetLeuSerProPro 974
|||||
3133 TCGGCAGCTCTACATTCAGGAAACCTTACAGCACATGCTCTCCCGCCCC 3182
|||||
974 InIleThrIleGlyProGlnArgProAlaProSerGlyProGlyGly 990
|||||
3183 AGATCACCATCGTCCCGAGAGGCGAGCTCCCTCAGGGCCAGGAGGGGC 3232
|||||
991 SerGlyGlyGlySerGlyGlyGlyGlyGlyGlyGlnProProProLeuG 1007
|||||
3233 AGTGTGGGGCAGTGGTGGGGCGGTGGGGCCAGCCACCTCCCTTGCA 3282
|||||
1007 nArgGlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLysProArg 1024
|||||
3283 GAGGGGCAAACTCAGCAGTTGACAGTGTAGTGTGCTGCCCAAGACCCGCG 3332
|||||
1024 roSerSerGlyAsnLeuLeuGlnSerProGluProSerTyGlyProAla 1040
|||||
3333 COTCCAGCGGGAACCTATTGTCAGTCCCGGGAACCAAGTTATGCTCTGCC 3382
|||||
1041 ArgProArgGlnGlnSerLeuSerLysGlySerIleGlySerGly 1057
|||||
3383 COTCCAGCGCACAGAGCTCAGCAAGAGGCGAGCATTTGGGGCGAGCGG 3432
|||||
1057 yGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyLeuLysProSerIleThrLys 1074
|||||
3433 GGGCAGCGGTGGCGAGGGGTGGGGGCTCAAGCCCTCCATCACCAGC 3482
|||||
1074 InHisSerGlnThrProSerThrLeuAsnProThrMetProAlaSerGlu 1090
|||||
3483 ASCATTCCAGACTCCATCCAGCTGAACCCAGCATGCGGCTCGGAG 3532
|||||
1091 ArgThrValAlaTrpValSerAsnMetProHisLeuSerAlaAspIleG 1107
|||||
3533 CGGACTGTAGCGTGGGTGTCGAATATGCCCTACCTGTCCGCTGACATCG 3582
|||||
1107 uSerAlaHisIleGluArgGluGluTyLysLeuLysGluTySerLys 1124
|||||
3583 GAGTGCACACATTGAGCGGGAAGAGTACAAGCTCAAGGAGTACTCGAAGT 3632
|||||
1124 erMetAspGluSerArgLeuAspArgValLysGluTyGluGluGlu 1140
|||||
3633 CCATGGACGAGCGCGCTGACAGAGGTGAAGGATGACGAGGAGATC 3682
|||||
1141 HisSerLeuLysGluArgLeuHisMetSerAsnArgLysLeuGluGlu 1157
|||||
3683 CACTCAGTGAAGGAAGGCTACATGTCCAAACCGGAAGTGAAGAGTA 3732
|||||
1157 rGluArgArgLeuLeuSerGlnGluGluGlnThrSerLysIleLeuMetC 1174
|||||
3733 CGAGCGGAGGCTGCTGTCCAGGAAGAGCAGACCAGCATCCTGATGC 3782
|||||
1174 InTyTGlnAlaArgLeuGluGlnSerGluLysArgLeuArgGlnGln 1190
|||||
3783 AGTACCAAGCCCGCTGGAGCAGAGCAAGCGCTTGAGGCGAGCAGCAG 3832
|||||
1191 ValGluLysAspSerGlnIleLysSerIleGlyArgLeuMetLeuVa 1207
|||||
3833 GTGGAGAAGGACTCCAGATCAAGAGCATCATTTGGCAGGCTGATGCTGGT 3882
|||||
1207 lGluGluGluLeuArgArgAspHisProAlaMetAlaAlaGluProLeuProG 1224
|||||
3883 GAGGAGGAGCTGCGCCGGGACCCCGCATGGCTGAGCCGCTGCCTG 3932
|||||
1224 luProLysLysArgLeuLeuAspAlaGlnArgGlySerPheProProTrp 1240
|||||
3933 AACCCCAAGAGGCTGCTCGACGCTCAGAGAGCGAGCTTCCCGCCCTG 3982
|||||
1241 ValGlnGlnThrArgVal...ArgTrpProHisLeuGlyThrAlaTrpPr 1256
|||||
3983 GTCCAACAACCCCGGTGTGACGCTGCGCCACCTTGGACGCGCTGGCC 4032
|||||
1256 oProGlnProHisProProHisProGlyCys 1266
|||||
4033 CCCCCAGCCCCACCCCCCCCCCCCCCGGTGC 4063
|||||
seq_name: gb_ro:AB016962
seq_documentation_block:
LOCUS AB016962 4801 bp mRNA ROD 01-SEP-2000
DEFINITION Rattus norvegicus mRNA for synGAP-d, complete cds.
ACCESSION AB016962
VERSION AB016962.1 GI:4417206
KEYWORDS synGAP-bi.
SOURCE Rattus norvegicus (sub_species:Sprague Dawley) cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 4801)
AUTHORS Suzuki, T.
TITLE SynGAP-d.
JOURNAL Published Only in DataBase (1999) In press
REFERENCE 2 (bases 1 to 4801)
AUTHORS Suzuki, T.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-1998) to the DBJ/EMBL/GenBank databases. Tatsuo Suzuki, Shinsu University School of Medicine, Department of Neuropathology; 3-1-1 Asahi, Matsumoto, Nagano 390-8621, Japan (E-mail:suzukit@sch.md.shinshu-u.ac.jp, Tel:+81-263-37-2683, Fax:+81-263-37-2725)
COMMENT On Mar 16, 1999 this sequence version replaced gi:4239945.
SEQUENCE Sequence updated (09-Feb-1999).
FEATURES
source
1..4801
/organism="Rattus norvegicus"
/sub_species="Sprague Dawley"
/db_xref="taxon:10116"
1124..4624
/genes="synGAP-d"
1124..4624
/genes="synGAP-d"
/notes="synaptic ras GTPase activating protein"
/codon_start=1
/product="SynGAP-d"
/protein_id="BAA74972.1"
/db_xref="GI:4417207"
gene
CDS
/translation="MEYFPAPFPFSGFLSRRLKSKTKSQPKLDRTSFSPQILP
RFSADHRLMQSFKSHSHSLSPSAAALELNLEDSIKVPVHSSILGQFEC
FEVTSSTGKCFACRSAAERDKWLENLQRAVKPNKNSRRVDNVLKLIIEARELPPK
KRYVCELDLDMLYARTTSKPSASGDTVFWEHGFENLPAVRLALHLRYRSDKKR
KKDKAGYVGLVTPVATLAGRHFTQWYPVTLPTGSGSGGSGSGSGSGSGSGGK"

```

KGCPAVRLKARYQTMIIIPMELYKEPARYVTNHYRMCLCAVLEPALNVKCKEEVASL
 VHILQSTGKAKDFSLDMAMSEVDREMERHLLIFRENTLATRAIEYMRLLGQKYLKDA
 IGEFIRALYSEENCVEVDPIKCTASSLAHQAOLRMCEALACKVVAHSHCVFPELKE
 VFASWRLCARERREDIADRLISASLFRFLCPALMSPLSLGMOEYPDQOTSRITL
 IAKVIONLANSFKTSKEDFLGFMEFLGLEWGSQQFLYEISNLDLTNSSSPGYI
 DLGELSTHALLAEVLPOLSKALLKGLPLRLLSDISTALRNPNQROPSROSER
 RSQPMVLGRPSAEMOGYMMRDLSNLSIDLOSPMARGLNSMDMARLPSTPEKPPPPP
 GGGKDLFYVRPPLARSSPAYCTSSSDITEPEQXMLSYNSYSMLDLOGDGPGRLLS
 SSVSNLAAGDLHSSQALTAALGURPAPAGRLKSGSGSSITAGMRLLSGMVTIDG
 VPAQRLIFLUSFNQPLFMAADGPPGAPGAGSGSGGPPSHHHHHHHHHHRGGEPPG
 DTPAFRIGYSKSDLSLSTGVKPPAASILHSHYSDFPGSGTDFTRQLSLQDNLQHM
 LSPQITIGPQRAPSPGGSGGGGGGGQPPQLQKSKQOLTVSAOKPRPSSG
 NLLQSPSPYGPAPROQSLSGSGSGSGSGGGGLKPSITKOHSONPSPINPT
 MPASERTVAVSNMHLISADIESAHIREEYKLYKEYSKMDESRLDRYERIEHSLKE
 PLHMSNPKLEYERLLSQEQTSKILMQYQARLRQSEKRLRQQQVEKDSQIKSIIS
 PSLOADAGGGGAAPGPPRG"

BASE COUNT 1043 a 1477 c 1314 g 967 t

alignment_scores:

Quality: 6308.50 Length: 1229
 Ratio: 5.179 Gaps: 5
 Percent Similarity: 99.105 Percent Identity: 99.024

alignment_block:

US-09-294-298-4 x AB016962 ..

Align seg 1/1 to: AB016962 from: 1 to: 4801

56 MetGluAlaLaProAlaLaProPheArgProSerGlnGlyPheLeuSe 72
 |||||
 1124 ATGAGTATTTTCCCGCTCGGCCCTTCGCAAGCTTCCCTGAG 1173
 72 rArgArgLeuLysSerSeriLeuArgThrLysSerGlnProLysLeuA 89
 |||||
 1174 CCGGAGGCTAAAGCTCATCAAGCTACAAAGTCAACACCAACTTG 1223
 89 spArgThrSerPheArgGlnIleLeuProArgPheArgSerAlaAsp 105
 |||||
 1224 ACCGGACACAGCTTTCGACAGATCTCGCTCCGCTCCGAGGTGTCAC 1273
 106 HisAspArgAlaArgLeuMetGlnSerPheLysGluSerHisSerHisG 122
 |||||
 1274 CATGACGGCCCGCGGTGTGTCAGAGCTTCAAGGAGTCTCACTCCCATGA 1323
 122 uSerLeuLeuSerProSerSerAlaAlaGluAlaLeuGluLeuAsnLeuA 139
 |||||
 1324 GTCCCTGCTGAGTCCCGCAGCAGTGTCTGAGGCCCTCGAGCTCAACCTGG 1373
 139 spGluAspSeriIleLysProValHisSerSeriLeuGlyGlnGlu 155
 |||||
 1374 ATGAAGACTCCATTATCAAGCCAGTACACAGCTCCATCCTGGGCCAGGAG 1423
 156 PheCysPheGluValThrThrSerSerGlyThrLysCysPheAlaCysAr 172
 |||||
 1424 TTCTGCTTTGAGTAAACAATCGTCTGGGACAAAATGTTTGGCTCTCG 1473
 172 qSerAlaAlaGluArgAspLysTrpIleGluAsnLeuGlnArgAlaVal 189
 |||||
 1474 GTCTGCACGCCAAGGACAAATGGATTGAGAACTACAGAGGGCTGTGA 1523
 189 ysProAsnLysAspAsnSerArgArgValAspAsnValLeuLysLeuTrp 205
 |||||
 1524 AACCAACAAGGACACAGCCGGGTAGATAACGTCTGAAACTATGG 1573
 206 IleIleGluAlaArgGluLeuProProLysLysArgTyrTyrCysGluLe 222
 |||||
 1574 ATCATAGAAGCTCGAGAGCTGCCCCCAAGAGCGATATTACTGCGAGTT 1623
 222 uCysLeuAspMetLeuTyrAlaArgThrThrSerLysProArgSerA 239
 |||||
 1624 ATGCTGGAGGACATGCTTATGACAGGACCTTCCAAAGCCCGCTCAG 1673

239 laSerGlyAspThrValPheTrpGlyGluHisPheGluPheAsnAsnLeu 255
 |||||
 1674 CCTCAGAGACACTGCTCTTTTGGGCGAGCAGCTTCGAGTTTAAACAACCTG 1723
 256 proAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysL 272
 |||||
 1724 CTGCTGCTCGGGCGCTGGCGGTGCATCTGTACCTGACTCGGACAAAA 1773
 272 sArgLysLysAspLysAlaGlyTyrValGlyLeuValThrValProValA 289
 |||||
 1774 GCGGAAGAAGGACAAAGCGAGGTACGTTGGCTGTGTGCTGCTCCAGTGG 1823
 289 laThrLeuAlaGlyArgHisPheThrGluGlnTrpTyrProValThrLeu 305
 |||||
 1824 CCACCTTGGCTGGCGGCCACTTTCACAGAGCAGCTGGTACCTCCCGTGA 1873
 306 ProThrGlySerGlyGlySerGlyMetGlySerGlyGlyGlyGlyGlyG 322
 |||||
 1874 CCAACAGGAAGTGGGGGCTCTGGGGGTATGGGCTCGGGGGAGGAGGGG 1923
 322 ySerGlyGlyGlySerGlyGlyLysGlyLysGlyGlyCysProAlaValA 339
 |||||
 1924 GTCAGGGGGCGCTCAGGGGGCAAGGAAAGAGGAGGCTGCTCTGCTGTC 1973
 339 rgLeuLysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeuTyr 355
 |||||
 1974 GGCTGAAGGCCGCTTACCAGACAATGAGTATCTGCCCATGGAGCTATAT 2023
 356 LysGluPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaVal 372
 |||||
 2024 AAGGAGTTTCAGANAATGTGACCAACCACTACCGCATCTGCTGCTGCT 2073
 372 lLeuGluProAlaLeuAsnValLysGlyLysGlyGluGluValAlaSerAla 389
 |||||
 2074 GCTGAGCCCGCCCTCAATGTCAAGGGCAAGGAGGAGGCTCGCTAGTGAC 2123
 389 euValHisIleLeuGlnInserThrGlyLysAlaLysAspPheLeuSerAs 405
 |||||
 2124 TGGTTTCATCCTGCAAGGACAGGCAAGGCAAGGCAAGGCAAGGCAAG 2173
 406 MetAlaMetSerGluValAspArgPheMetGluArgGluHisLeuIleph 422
 |||||
 2174 ATGGCCCATGTGAGAGGTAGACCGGTTCATGGAGCGGGGAAACACTCAT 2223
 422 eArgGluAsnThrLeuAlaThrLysAlaIleGluGluTyrMetArgLeuI 439
 |||||
 2224 CCGCGAGAACACGCTCGCCACTAAAGCCATAGAAGAGTATATGAGACTGA 2273
 439 leGlyGlnLysTyrLeuLysAspAlaIleGlyGluPheIleArgAlaLeu 455
 |||||
 2274 TTGGCCGAATAACCTCAAGGATGCCATTTGGGGAGTTTCATCCGGGCTG 2323
 456 TyrGluSerGluGluAsnCysGluValAspProIleLysCysThrAlase 472
 |||||
 2324 TATGAATCTGAGGAGAACTGTGAAGTAGACCCCATCAAGTGCACAGGCTC 2373
 472 rSerLeuAlaGluHisGlnAlaAsnLeuArgMetCysGlyLeuAlaAla 489
 |||||
 2374 CAGTCTGGCAGAGCACCCAGGCCAACCTCGCGATGTGCTGTGAGTTGGCC 2423
 489 euCysLysValValAspSerHisCysValPheProArgGluLeuLysGlu 505
 |||||
 2424 TGTGCAGAGGTGGTCAACTCCCATTTGGTGTTCAGGAGGAGCTGAAGGAG 2473
 506 ValPheAlaSerTrpArgLeuArgCysAlaGluArgGlyArgGluAspI 522
 |||||
 2474 GTGTTTGCATCATGGCGCTGCGCTGTGTCAGAGCGGGCGCGGAGGACAT 2523
 522 eAlaAspArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProA 539
 |||||
 2524 TGCTGACAGGCTGATCAGCGCTTCGCTCTTCCTTCGCTTCTCTGCCCCG 2573

539 laLeMetSerProSerLeuPheGlyLeuMetGlnGluTyrProAspGlu 555
|||||
2574 CCATCATGTGCGCCAGCTCTGTTGGAGCTGATCAGAGGATACCCAGATGAG 2623
556 GlnThrSerArgThrLeuThrLeuLeuAlaLysValIleGlnAsnLeuAl 572
|||||
2624 CAGACCTCAGAACCTCACCCTCATCGCAAGGTTATCCAGAACCTGGC 2673
572 aaanPheSerLysPheThrSerLysGluAspPheLeuGlyPheMetAsnG 589
2674 CAACCTTTTCAAGTTTACCTCAAGAGGAGGACTTCTTGGGCTTCATGAACG 2723
589 luPheLeuGluLeuGluTTPGlySerMetGlnGlnPheLeuTyrGluIle 605
2724 AGTTTCTGGAGCTGGAGTGGGTTCTATGCACCAATTCCTGTATGAGATA 2773
606 SerAsnLeuAspThrLeuThrAsnSerSerSerPheGluGlyTyrIleAs 622
2774 TCCAACCTTGACACACTGACCAACAGCAGCAGTTTTTGAGGGCTACATAGA 2823
622 pLeuGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuP 639
2824 CTTGGGCGGAGCTCTCACACTTCACGCCCTGCTCTGGAGGTGCTGC 2873
639 roGlnLeuSerLysGluAlaLeuLeuLysLeuGlyProLeuProArgLeu 655
2874 CCAGCTTCAGCAAGAGGCCCTCCTGAAGCTGGGCCGCTGCCCGGGCTC 2923
656 LeuSerAspIleSerThrAlaLeuArgAsnProAsnIleGlnArgGlnPr 672
2924 CTCAGCGACATCAGCACAGCCCTCAGGAACCTTAACATCCAAAGCGACCC 2973
672 oSerArgGlnSerGluArgAlaArgSerGlnProMetValLeuArgGlyP 689
2974 GAGCGCGCAGAGGAGCGCGCTCGCTCTCAGCCCATGTTGCTGCGCGGGC 3023
689 roSerAlaGluMetGlnGlyTyrMetMetArgAspLeuAsnSerSerIle 705
3024 CGTCAGCGAGATGACGGCTACATGATGCGGGACCTCAACAGCTCCATC 3073
706 AspLeuGlnSerPheMetAlaArgGlyLeuAsnSerSerMetAspMetAl 722
3074 GACCTTCAGTCTTCATGCTCGAGGCCCTCAACAGCTCTATGGACATGGC 3123
722 aaArgLeuProSerProThrLysGluLysProProProProProGlyG 739
3124 TCSCCTCCCTCCCAACCAAGAGAGAACCCCGCGCCCTCCCGGGTG 3173
739 lyGlyLysAspLeuPheTyrValSerArgProProLeuAlaArgSerSer 755
3174 GGGTAAAGACCTGTTCTATGTAGCGCGGCCACCACTGGCGCGGTCTCC 3223
756 ProAlaTyrCysThrSerSerSerAspIleThrGluProGluGlnLysMe 772
3224 CCAGCACTACTGCACGAGCAGCTCGGACATCACAGAGCCGGAGCAGAGAT 3273
772 tLeuSerValAsnLysSerValSerMetLeuAspLeuGlnGlyAspGlyP 789
3274 GCTGAGTGTCAACAGAGTGTCTCATGCTGGACCTGCAGGGCGACGGGC 3323
789 roGlyGlyArgLeuAsnSerSerValSerAsnLeuAlaAlaValGly 805
3324 CTGGGGCGCGCTTAAACAGCAGTAGTGTTCACACTGGCAGCTGTTGGG 3373
806 AspLeuLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuAr 822
3374 GACCTGTGCACTCAAGCAGGCTTCACTGACAGCAGCCTTGGGGTTGGC 3423
822 gProAlaProAlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAr 839
3424 GCCTGACCTGCGGGCGGCTCTCCCAAGGAGTGGCTCTTCCATCACAG 3473
839 laAlaGlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValPro 855

3474 CAGCGCGCATGCGCTCAGCGAGATGGTGTCTACTACGGATGGTGTCCCC 3523
856 AlaGlnGlnLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMe 872
3524 GCCCAGCAGCTGCGCATCTCTTCTTCCCTCCAGAACCTCTCTTCCATAT 3573
872 taAlaAlaAspGlyProGlyProProAlaGlyHisGlyGlySerSerGlyH 889
3574 GGCTGCCGATGGACAGCGGCCCGCCAGCAGGCCATGGAGGAGCAGTGGCC 3623
889 lsGlyProProSerSerHisHisHisHisHisHisHisHisHisHisArq 905
3624 ATGGTCCACCTTCTCTCCATCACCACCACCACCACCACCATCACCATCACC 3673
906 GlyGlyGluProProGlyAspThrPheAlaProPheHisGlyTyrSerLy 922
3674 GGGGAGAAACCCCGAGGGGACACTTTTGGCCCCGTTCATGGCTATAGCAA 3723
922 sSerGluAspLeuSerThrGlyValProLysProProAlaAlaSerIleI 939
3724 GAGCGAGGACCTCTCTACAGGGTCCCTAAGCCCCCTCGGCCCTCCATCC 3773
939 euHisSerHisSerTyrSerAspGluPheGlyProSerGlyThrAspPhe 955
3774 TTCAGCGCACAGCTACAGTGATGAGTTTGGAGCCCTCTGGTACTGATTTT 3823
956 ThrArgArgGlnLeuSerLeuGlnAspLeuGlnHisMetLeuSerPr 972
3824 ACCCGTCGGCAGCTCTCATTACAGGACCACTACAGCACATGCTCTCCCC 3873
972 oProGlnIleThrIleGlyProGlnArgProAlaProSerGlyProGlyG 989
3874 GCCCGAGATCACCATCGTCCCCAGAGGCCAGCTCCCTCAGGGCCAGGAG 3923
989 lyGlySerGlyGlyGlySerGlyGlyGlyGlyGlyGlyGlnProProPro 1005
3924 GGGCAGTGGTGGGGCAGTGGTGGGGCGGTGGGGCCAGCCACCTCCC 3973
1006 LeuGlnArgGlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLysPr 1022
3974 TTGCGAGGGGGCAATCTCAGCAGTTGACAGTGGTGGTGGTGGTGGTGG 4023
1022 oArgProSerSerGlyAsnLeuLeuGlnSerProGluProSerTyrGlyP 1039
4024 CCGGCGCTCCAGCGGGAACCTATTGACAGTCCCGGAACCAAGTTATGGTC 4073
1039 roAlaArgProArgGlnGlnSerLeuSerLysGluGlySerIleGlyGly 1055
4074 CTGCCCCGCCAGCGCAACAGGCTCAGCAAGAGGGCAGCATTTGGGGC 4123
1056 SerGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyLeuLysProSerIleTh 1072
4124 AGCGGGGCGAGCGGTGGCGGGGGTGGGGGCTCAAGCCCTCCCATCAC 4173
1072 rLysGlnHisSerGlnThrProSerThrLeuAsnProThrMetProAlaAs 1089
4174 CAAGCAGATTCCCGAGCTCCATCCAGCTGACCTGAACCCACGATCCCGGCT 4223
1089 erGluArgThrValAlaTrpValSerAsnMetProHisLeuSerAlaAsp 1105
4224 CGGAGCGGACTGTAGCTGGGTGCTCAATATGCTCCTCACCTGCTCCGCTGAC 4273
1106 tLeuLysSerAlaHisIleGluArgGluGluTyrLysLeuLysGluTyrSe 1122
4274 ATCGAGAGTGCACACATTGAGCGGGAAGAGTCAAGCTGAAGGAGTACTC 4323
1122 rLysSerMetAspGluSerArgLeuAspArgValLysGluTyrGluGluG 1139
4324 GAAGTCCATGGACGAGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4367
1139 luIleHisSerLeuLysGluArgLeuHisMetSerAsnArgLysLeuGlu 1155

```

4368 AGATCCACTACTGAAGAAAGGCTACACATGTCCACCGGAAGCTGGAA 4417
1156 GlnTyrGluArgArgLeuLeuSerGlnGluGlnThrSerLysIleLe 1172
|||||
4418 GAGTACAGCGGAGCGCTGTCTCCAGAGAGAGACAGACAGCAGATCCT 4467
1172 uMetGlnTyrGlnAlaArgLeuGluGlnSerGlnLysArgLeuArgGln 1189
|||||
4468 GATGACGTACCAAGCCCGCTTGAGGAGAGCGGAGAGCGCTTGAGGAGC 4517
1189 GlnValGluLysAspSerGlnIleLysSerIleIleGly..... 1202
4518 AGCAGGTGGAGAGACTCCAGATCAAGAGCATCTTGGCAGCCCGTCC 4567
1203 ...Arg..LeuMetLeuValGluGluGlnLeuArgArgAspHisProAlaM 1218
|||||
4568 CTTGAGGCTGATGTGTGTGAGAGAGAGTGGCGCGGAGACACCCGCCA 4617
1218 eTAlaGluProLeuProGluProLysLysArgLeuLeuAspAlaGln.Ar 1234
|||||
4618 TGGCTGAGCGCTGCTGCTGAACCAAGAGAGCTGCTGAGCGCTCAGGAG 4667
1234 gClySerPheProProTyrValGlnGlnThrArgVal...ArgTyrProH 1250
|||||
4668 AGGCAGCTTCCCGCTTGGGTCCAAACCAACCGCTGTGACGCTGGCCCC 4717
1250 IsLeuGlyThrAlaTyrProGlnProHisProHisProHisProGlyCys 1266
|||||
4718 ACCTTGAAGCGCTGCGCCCCCCCCAGCCGCCCCCCCCCCCGCGCTGC 4767
1267 ArgSerGlnArgThrAlaSerSerGlyThrPro 1277
4768 AGATCACAGAGAACGGCGAGTTCGGGAACACCG 4800

seq_name: gb_htg20:AL161903

seq_documentation_block:
LOCUS AL161903 150956 bp DNA HTG 29-SEP-2000
DEFINITION Homo sapiens chromosome 6 clone RP11-175A4, *** SEQUENCING IN
PROGRESS ***, 3 unordered pieces.
ACCESSION AL161903
VERSION AL161903.12 GI:10039702
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tracy,A.
Direct Submission
Submitted (28-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerquest@sanger.ac.uk
On Sep 9, 2000 this sequence version replaced gi:9796876.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA175A4
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 99% of reads
Chemistry: Dye-primer-amersham; 0% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 145680 bases at least Q40
Consensus quality: 147360 bases at least Q30
Consensus quality: 148600 bases at least Q20
Insert size: 150756; sum-of-contigs
Insert size: 148326; 10.8% error; agarose-fp
Quality coverage: 6.43x in Q20 bases; sum-of-contigs Quality

```

```

coverage: 6.54x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 54611: contig of 54611 bp in length
* 54612 54711: gap of 100 bp
* 54712 70293: contig of 15582 bp in length
* 70294 70393: gap of 100 bp
* 70394 150956: contig of 80563 bp in length.
FEATURES
source
1..150956
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-175A4"
/clone_lib="RPCI-11.1"
misc_feature
1..54611
/note="assembly_fragment:01454"
vector_end:T7
vector_side:left
misc_feature
54712..70293
/note="assembly_fragment:02542"
fragment_chain:1"
70394..150956
/note="assembly_fragment:02844"
fragment_chain:1
vector_end:SP6
vector_side:right"
BASE COUNT 39575 a 37381 c 36411 g 37389 t 200 others
ORIGIN

alignment_scores:
Quality: 4066.00 Length: 3227
Ratio: 3.836 Gaps: 22
Percent Similarity: 32.848 Percent Identity: 31.887

alignment_block:
US-09-294-298-4 x AL161903/rev ..
Align seg 1/1 to reverse of: AL161903 from: 1 to: 150956
46 ArgValSerProGlyLysGlnTyrSerMetGluAlaAla.ProAlaA 62
|||||
107790 CGCTTGAGCCAGGAGCGAGAGTTGCGACGAGCGAGATCACACCACTG 107741
62 la.....ProPheArgProSerGlnGlyPheLeuSerArgArgLeuLys 76
:: ||| ||| |||:: ::|||::: |||
107740 TACTCCAGCTGGCGGACAGAGCAAG...ACTGCTCAAAAAAAAAAAAA 107694
77 SerSerIleLysArgThrLysSerGlnProLysLeuAspArgThrSerSe 93
::: ||| |||::: |||
107693 AAAAAAAAAAAGAGAGCTGGGAGCCCC.....TGATCCCTTCC 107653
93 rPheArgGlnIleLeuProArgPheArgSerAlaAspHisaspArgAlaA 110
::: ||| |||
107652 TCTCTGAACCTCTCGCTGCC.....AGGGCCC 107624
110 rgLeuMetGlnSerPheLysGluSerHisSerHisGluSerLeuLeuSer 126
|||||
107623 GGCTGATGCAAGCTTFAAGGAGTCACTCTCATGAGTCTTGCTGAGT 107574
127 ProSerSerAlaAlaGluAlaLeuGluLeuAsnLeuAspGluAspSerIl 143
|||||
107573 CTTAGCATGAGCTGAGCATTTGAGCTCACTTGGATGAAGATTCAT 107524
143 eIleLysProValHisSerSerIleLeuGlnGlnPheCysPheGlu. 159

```

```

|||||
107523 TATCAAGCCAGTGCACAGCTCCATCTCGGGCCAGGAGTCTCTTTTGGG 107474
160 .val..... 160
:::
107473 TACTGGGTCTGGTGGGTGGGAGGCCAAAGACAGGGGTGATGGAAGG 107424
160 ..... 160
107423 TGGGGGCGAGAGAGTCTAGAGAAAGTGGCACAGGTGGGACATCAAGGA 107374
160 ..... 160
107373 AACAGAACTTCTGGGACTGGAGAAAGGTAGGCCAGGGAGGAAGAGAA 107324
160 ..... 160
107323 GGTAGCAGAGTCTCTCCCTCTGTAGCCCTTTCCCTCAACTCCACACTC 107274
161 .....ThrThrSerSerGlyThrLysCysPheAlaCysArgSer 173
|||||
107273 CTTCCTAGGTAACAACTTCATCAGGAACAAATGCTTTGCCCTGTCGGTCT 107224
174 AlaAlaGluArgAspLysTrpIleGluAsnLeuGlnArgAlaValLysPr 190
|||||
107223 GCGGCCGAAGAGACAAATGATTTGAGATCTGCAGCGGGCAGTAAAGCC 107174
190 oAsnLys..... 192
|||||
107173 CAACAAGTATTGGGAATAAAGGGGACACAACTCTGTAGGGCAAGGT 107124
192 ..... 192
107123 TCCACCAACACGGGAAGCTGGGGCCCTAGGGAGGAAGTGAAGTAAAGGA 107074
192 ..... 192
107073 GGAGAGGCTTGGGAAGGAGAGGATTTAGGTACAGTGTATCTGGACAAGC 107024
192 ..... 192
107023 AGGGGGAGACCCCATTTATCTGAGTCCCCCAATTTCTTTTCGCTTCGT 106974
192 ..... 192
106973 ACTGCTACCCCTTACGATCTCTTCCCTGCCATAGACATAGACT 106924
192 ..... 192
106923 TACAGAGTTGATGGGGCTCTGGAATTTCTTAGTCTAGCTTCCCTGCAGG 106874
192 ..... 192
106873 CAGGAATCTTCACTAATACCCCGCAGGACACATCAATACACTTAGCCA 106824
192 ..... 192
106823 AGTTTCTGACTTGGTTTCCTCTCTAAGACAGAGGGAACCTGGTGGCAAT 106774
192 ..... 192
106773 GGGTTGGATTAAATGATCTCTAGGTCCTTTGGCACATAAATTTCTATGA 106724
192 ..... 192
106723 GTCTGTCTTCCAAAATTCATGCTTCCAGTTCTGTCTGACATATTCCTG 106674
192 ..... 192
106673 TATCCAGTTTCTAGCAAGCTTACAGTCTTACAGTCTTACACTCCCTCTGGGGG 106624
192 ..... 192

```

```

106623 AACTCTTGTGTCTTGATGCTCTTAAAGAACTCAACCCAAAGATGCTGACAT 106574
192 ..... 192
106573 GATCTGACCTATGACAGAGTACAACACCATGTTCCCTTTTCAACACACTGC 106524
192 ..... 192
106523 AATATCACTGAGCTTACCATTAGCTTGTGCTCAATTTGTATCCCCCAG 106474
192 ..... 192
106473 GCCCTTACTGATTTCTTTTCTTTTGTGAGACAGAGTCTCACTCTG 106424
192 ..... 192
106423 TCACCCAGGCTGGAGTGCAGTGGCGTGATCTTGGCTCACTGCAACCTCG 106374
192 ..... 192
106373 CCTCTGGGTTTAAAGCAATTCCCTGTGCTCAGCTCCCCCAGTAGTGGA 106324
192 ..... 192
106323 TTACAGGCATGCACACACACACCGACTAATTTTGTATTTTAGTAGAG 106274
192 ..... 192
106273 ACGAGTTTTCACCATGTTGGCCAGGCTAGTCTCAAACTCTGACCTCAG 106224
192 ..... 192
106223 TTGCTGTGCTGCCTCAGCTCCCAAGTGTAGGATTACAGGCATGAGCC 106174
192 ..... 192
106173 ACCACATCTGCCCAAGCTTTACCAATCTATATGCAATCTTTTTCACA 106124
192 ..... 192
106123 ATTTCTGATAGTCTCTGCAGGACTTCCAGTTCCTTCCAAATTTCTACAT 106074
192 ..... 192
106073 AATATTTTGGCTGTATTTCCAGCTTTTGAATCTTTCCAATACTGTTT 106024
192 ..... 192
106023 GTGTTGTAGTGTATGTCTAGCTTCTACAGTAGCTTCCCTGGGGCTGCT 105974
192 ..... 192
105973 ATAACAACACTGCTACAAACTGGGTCACTAAACACAGAAATTTATTCT 105924
192 ..... 192
105923 CTCACAGTTCAGGAGGCTGGAAGTCCAAAAGCAAGGTATACAGAGGCCA 105874
192 ..... 192
105873 CGCTCTCTTGTGCTGTAGCAGGGAATCTCTCTTGGCTTCTCCTAGCT 105824
192 ..... 192
105823 TCCGAGGTTGCCAGAGTCTTGGCAATCTCTGGGCTTATAACTGCATCC 105774
192 ..... 192
105773 GTCTAATTTCTGCCTCCATCTTCATGTAGCTGGGCTCCTTCTGTGTCT 105724
192 ..... 192
105723 CTGTATCTGTATCTCTGTGCTCCAAATCTCCCTCTCCATATAAGACA 105674

```

192 192
105673 CCAGTTTTTATAAGGTGGGTTAAGGTCACCTCTAATTCAGTATGGCCTC 105624
192 192
105623 TTTTTTTTGAGATGGAATATCGCTCTTTGTTCCAGGCTTGGAGTGCAAT 105574
192 192
105573 GGCATGAGTCAGCTCAGCTCAGACCTCCGCTTCCAGGTTCAAGTGATTC 105524
192 192
105523 TCCTGCCTCAGCCTCCCTGAGTACGTGGGATACAGGCACGTGCCACTATG 105474
192 192
105473 CCCAGTAAATTTGTTTGTATTTTATTATAGACGGGTTTCACTATGT 105424
192 192
105423 TGGCCAGCTGGCTCGAATCTCTGACCTCAGGCAATCCACCCTCGGC 105374
192 192
105373 CTGCCAAAGTCTGGGATTACAGGAATGAGATACCATGGCTGGCCTCTTC 105324
192 192
105323 TTAGTTGATTACATTTGTCAGGATCCTGTTTCCAAATTAAGTACATTC 105274
192 192
105273 ACAGTTTCAGGTTAGACATGAGTTTGTGGTACTATTCAACCCAGTA 105224
192 192
105223 CACCGCTCACCTTATGATTGCTGCTGTATCCCACCTTCTTACGTGCTC 105174
192 192
105173 TCCTCTTTTGGTGTCTCTGTCTCCCTTCCCTTAAACATCATGCCACC 105124
193 205
105123 CCACCATGCCAGGACACACGCCCGGGTAGACAATGTGCTAAAGCTGTG 105074
205 pIleileGluAlaArgGluLeuProLysLysArgTyrTyrCysGluL 222
105073 GATCATAGGCGCGGAGCTGCCCGCCAGAGCGGTACTACTGTGAGC 105024
222 euCysLeuAspMetLeuTyrAlaArgThrThrSerLysProArgSer 238
105023 TCTGCTGGATGACATCTGTATGCACGCCACCCTCCAAAGCCGCTCT 104974
239 AlaSerGlyAspThrValPheTgGlyGluHisPheGluPheAsnLe 255
104973 GCCTCTGGGACACCGTCTCTGGGGGAGCACTTCGAGTTTAAACACT 104924
255 uProAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysL 272
104923 GCCGCTCTCGTCCCTCGGGGTGCATCTGTACCGTACTCAGACAAAA 104874
272 ysArgLysAspLysAlaGlyTyrValGlyLeuValThrValProVal 288
104873 ACCGCAAGACGACAGGAGGATATGTGGCCTGTGTGCTGTGCCAGTG 104824
289 AlaThrLeuAlaGlyArgHisPheThrGluGlnTrpTyrProValThrLe 305
104823 GCCACCTGGCTGGCGCCACTTCACAGACAGTGTACCTGTAAACCT 104774

305 uProThrGlySerGlyGlySerGlyMetGlySerGlyGlyGlyGlyG 322
104773 GCCAACAGGCAGTGGGGATCTGGGGCATGGGTTCCGGAGGGGAGGG 104724
322 lysSerGlyGlyGlySerGlyLysGlyLysGlyLysGlyCysProAlaVal 338
104723 GCTCGGGGTGGCTCAGGGGCAAGGCAAGAGGTTGCCCGCTGTG 104674
339 ArgLeuLysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeuTy 355
104673 CGGTGAAGACAGCTTACCAGACAATGAGCATCTTGGCCATGGAGCTATA 104624
355 rLysGluPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaV 372
104623 TAAAGGTTTGCAGAGTATGTCACCAACCATTTATCGATGCTGTGTGCG 104574
372 alLeuGluProAlaLeuAsnValLysGlyLysGluGluValAlaSerAla 388
104573 TCTTGGAGCCCGCTGAATGTCAAAGGCAAGGAGGAGGTTGCCAGTGCA 104524
389 LeuValHisIleLeuGlnSerThrGlyLysAlaLys 400
104523 CTAGTTCAATCTCTGCAGAGTACAGCAAGGCCAAGGTGAGTGTGTGCC 104474
400 400
104473 CTCAGGAAAGGTGACTTGGGAATGGGCACCTTCTTGGGGGTTAGTGAG 104424
400 400
104423 ACAGGCCAAATTCAGGAGATTGGGTTGTCAGAGGCTGACACTTGGATT 104374
401 AspPheLeuSerAspMetAlaMetSerGluValAspA 413
104373 TCCTGGCCCTCAGGACTTCTTTCAGACATGGCCATGCTGAGGTAGACC 104324
413 rgPheMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAlaThr 429
104323 GGTTCATGGAACGGGAGCCTCATATTCGCGAGAACAACGCTTCCCACT 104274
430 LysAlaIleGluLutyrMetArgLeuIleGlyGlnLysTyrLeuLysAs 446
104273 AAAGCCATAGAAGATATATGAGACTGATTGGTCAGAAATACCTCAAGGA 104224
446 pAlaIle 448
104223 TGCCATTGGTATGGCCACACTCAGGCCCTCTTCTTCCAAACCTGCCAG 104174
448 448
104173 ATGTCCACCCCCAGACCCCAAGTCCACCTTCCACAGCTTGATACTCTA 104124
448 448
104123 ACCCAGAGTCTTAGGACTCCAGCCTCCACACCTGATTCGAAATTTCCC 104074
448 448
104073 CAACCTGGCCACCCCTTCCCTGGCTTGGAAAGTGTGACCACACCCCTC 104024
449 GlyGluPheIleArgAlaLeuTyrGluSerGl 459
104023 TTGTGCCCCACCCCCAGGAGAAATTCATCCGCTGCTCTGATGAATCTGA 103974
459 uGluAsnCysGluValAspProIleLysCysThrAlaSerSerLeuAlaG 476
103973 GGAAACTGCGAGGTAGACCTTATCAAGTCACAGCATCCAGTTTGGCAG 103924
476 luHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysLysVal 492
103923 AGCACAGGCCAACCTCGGAATGTGTGTGGCTTGGCCCTGTGCAAGGTG 103874
493 ValAsn 494

|||||
103873 GTCACACTCCCACTGGTGTGAGACTGGGAACGCTGGGCTGGGGGCCAGGGTC 103824
494 494
103823 GGGGAATTATGTCTTCATCTGTTCATCTATCTGTCCATCCCTCAAGAGG 103774
494 494
103773 ACTGAGCACCATTATATGGGCAAGCATTTGTCTAGCGCTATAGAGCAA 103724
494 494
103723 CAGGTGAAGAGGCTTGGTCCCTGCCCTCAGAGGGCTCCACCAGAATGG 103674
494 494
103673 GGACAAATTAGAGAAAAAAGGCAAGCCAGAGCCATTAATGGTGTG 103624
494 494
103623 TAAGTGCTGAGTAAGGCTGCCCAACCTCTGTGTGACATAAGTCAAG 103574
494 494
103573 AGAAGCAGAGCTTTGAGATAGTGGGAAGAGGTGCCCTTGGGTAGG 103524
494 494
103523 CTTTGAAGACTGGTTTAGTTCATATATGGACATAGTTGGCAAGAAAG 103474
494 494
103473 ACATTTCCAGAAGAGGCTGTGAGAAAGGCACATGTGTGATGCTGAAAGG 103424
494 494
103423 CCCAGGAGTTTTCAGGGACATTAAGTAGTTAGTACATATACATCAG 103374
494 494
103373 GTTTAGTGGAGCATGTCCCTCATATAATGGGAGTGGCGGAGAGATGTCTG 103324
494 494
103323 GGCAGGAAGATTAGCTTTAGAAAACCTGGAAGGCTCAAGGAGTCTGAGTC 103274
494 494
103273 ATTTGAGGCTTGGGATGCCATTAAAGGTGCAGAAATATTGTGATTTA 103224
495 SerHis_Cys 497
103223 GAAGATTAAATCCATAGCTGGGACGGTGGCTCACACCTGTAAATACCAGC 103174
497 497
103173 ACTTTGGAGGCTGAGCGGGCAGACCACTGAGTCAAGGAGTTTGAGAC 103124
497 497
103123 CAGGCTGACCAACATGGAGAAACCTGTCTCTACTAAAAATACAAAAATTA 103074
497 497
103073 GCCAGATGTGGTGGCACATGCCCTGTAAATCCCAGCTATTGGGAGGCTAAG 103024
497 497
103023 GCAGGACAATCACTTGAATCTGGGAGGTGGAGTTGTAGTGAGCGGAGAT 102974
497 497

102973 CACACCATTTGCACCTCTAGCCTGGGCAACAAGAGTGAAACTCCCTCTCAAA 102924
497 497
102923 AAAAAAAATTAATCCATCAAGTAGAGTGTATGTAGTAGTAGAGGAGATG 102874
497 497
102873 TATGTTGGGGATGACTGGGAATAGAGCCAGCTGGGGCTCGGAGGCATGA 102824
497 497
102823 AAGTCAGATCCTGAATCAGAACAGTGACAGAAGTTAGGAAGGAGCTACTG 102774
497 497
102773 GAAGGACCCTCTCGGAAAGGATCAGTAGGAATGTCAAGCTTATTGGAAA 102724
497 497
102723 GGAGGAGCATGTCTGTGGGAGTCAAGGATGACTCAAGAGGCATGAGG 102674
497 497
102673 CTGTGTGGTTGGAGACCGGTGGCCCTATTGACAACCAAGGAAAGTCAGGAG 102624
497 497
102623 GAGGAGCAGTTGGAGGTGTGGGGCGATGGTGAGCTCTGCTTTACACCA 102574
497 497
102573 GCCGAGTTTAAGGTCTCAGTGGGACATTTGAAGTGAACCTGGGAGGTGCGG 102524
497 497
102523 GGAGTGAGCTCTGAGCCATTCCAGGAGCTGGGATCATGCCCTGGGACCC 102474
497 497
102473 TCATATCCCAATTTCCCTGGAATCCAGAAGAGTTGGGGGTCAGAGCTCC 102424
497 497
102423 TGTACCTCAAGTGACCCCTCCATCTCTCTCCCATCTCTGCTCTCCCTGGT 102374
497 497
102373 GTCTGTTTTTCT 102324
497 497
102323 CTCTCTCCACGCTCTCTCTCCCTCACTCTCTCTCTCTCTCTCTCTCTCTCT 102274
497 497
102273 CTCCCTAAATCTCTGTTCCTCTGCGATGGCCCTTCTTCAAGCAGCC 102224
497 497
102223 TCCCATCTTGTCTCTGGGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 102174
497 497
102173 TCCACACCCCTCACCTTCCTACACCCCTCAGCATGTTCCTGGAAGCTG 102124
497 497
102123 AGGGTCTCTGGGCTCAGTCCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 102074
498 ValPheProArgGluLeuLysGluV 506
102073 TGTCTCCCCGACCCCTTCCCCCGACGGTGTTCCTCCGAGGAGCTGAAGGAGG 102024


```

repeat_region 4903..4970
repeat_region /note="34 copies 2 mer ga 70 conserved"
repeat_region 4996..5190
misc_feature /note="13 copies 15 mer 59 conserved"
misc_feature complement(6393..6846)
misc_feature /gene="dJ570F3.1"
misc_feature /note="match: GSS: Em:AQ475907"
repeat_region 8077..8178
repeat_region /note="MLR1C repeat: matches 6..109 of consensus"
repeat_region 8476..8841
repeat_region /note="MLR1C repeat: matches 109..461 of consensus"
repeat_region 10989..11119
repeat_region /note="FLAM_C repeat: matches 3..131 of consensus"
repeat_region 13130..13161
repeat_region /note="16 copies 2 mer ca 93 conserved"
repeat_region 13256..13303
repeat_region /note="24 copies 2 mer ct 75 conserved"
repeat_region 17111..17815
misc_feature /note="CpG island"
misc_feature /evidence=not_experimental
repeat_region 17369..17458
repeat_region /note="6 copies 15 mer 70 conserved"
repeat_region 17370..17513
repeat_region /note="72 copies 2 mer cc 63 conserved"
repeat_region 17515..17658
repeat_region /note="72 copies 2 mer gg 59 conserved"
repeat_region 17631..17726
repeat_region /note="32 copies 3 mer gcc 63 conserved"
repeat_region 19005..19078
repeat_region /note="37 copies 2 mer ga 67 conserved"
repeat_region 19612..19691
repeat_region /note="40 copies 2 mer cc 63 conserved"
repeat_region 19858..19911
repeat_region /note="27 copies 2 mer cc 70 conserved"
repeat_region 24653..24808
repeat_region /note="78 copies 2 mer cc 57 conserved"
repeat_region 24706..24831
repeat_region /note="14 copies 9 mer ctctctct 61 conserved"
repeat_region 25104..25289
repeat_region /note="93 copies 2 mer cc 59 conserved"
repeat_region 25187..25294
repeat_region /note="12 copies 9 mer ctcccccc 64 conserved"
repeat_region 25367..25465
repeat_region /note="11 copies 9 mer gggagagag 69 conserved"
repeat_region 25373..25522
repeat_region /note="10 copies 15 mer 64 conserved"
repeat_region 25413..25469
repeat_region /note="19 copies 3 mer agg 86 conserved"
misc_feature 26820..27191
misc_feature /note="match: GSS: Em:AQ605113"
misc_feature 27145..27728
misc_feature /note="CpG island"
misc_feature /evidence=not_experimental
misc_feature complement(27412..27910)
misc_feature /note="match: GSS: Em:AQ369330"
misc_feature join(27713..28149,28320..28379,28488..28533,28661..28710,28854..29088)
misc_feature /gene="C1CK0721Q.5"
misc_feature /note="match: cDNAs: Em:AF106943 Em:E06811
misc_feature match: ESTs: Em:W75983 Em:AA187766 Em:AI004315 Em:R77008
misc_feature Em:N94792 Em:AA843316 Em:AA384803 Em:AI245618 Em:AI921804
misc_feature Em:AI792323 Em:AI802586 Em:AI524460 Em:AA843230
misc_feature Em:AA897246 Em:AI096824 Em:AI929640 Em:N32014 Em:AA725127
misc_feature Em:AA357305 Em:AA218432 Em:AI332801 Em:AA090206
misc_feature Em:AI050991 Em:AF034177 Em:W39055 Em:AA622278 Em:AA069440
misc_feature Em:AI087362 Em:AI168128 Em:AI340323 Em:AI028045
misc_feature Em:AA934024 Em:AI073391 Em:AA331045 Em:AA708233
misc_feature Em:AI342398 Em:AA056366 Em:AI582796 Em:AI762297
misc_feature Em:AI340198 Em:AI541387 Em:N34313 Em:AI567726 Em:AI364919
misc_feature Em:AI310946 Em:AA181976 Em:AI042256 Em:R81592 Em:T49588
misc_feature Em:AI929571 Em:D52353 Em:AA035663 Em:N48534 Em:AI373814
misc_feature Em:AA971499 Em:H25294 Em:AA703121 Em:AA46226 Em:AA379516
misc_feature Em:AI133425 Em:W70287 Em:AA937820 Em:H21228 Em:AA471384

```

```

Em:N71594 Em:AA059718 Em:AA377858 Em:H25235 Em:H52756
Em:D52710 Em:AA329903"
/evidence=not_experimental
/product="dJ570F3.2 (POLYPEPTIDE FROM PATENTED CDNA
EMBL:E06811)"

alignment_scores:
  Quality: 4060.50      Length: 3227
  Ratio: 3.838          Gaps: 21
  Percent Similarity: 32.786 Percent Identity: 31.856

alignment_block:
US-09-294-298-4 x HSDJ570F3/rev ..
Align seg 1/1 to reverse of: HSDJ570F3 from: 1 to: 94770

46 ArgValSerProGlyGly.....LysGlnTyrSerMetGluAla 58
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10638 CGCTTGAGCCCGAGGAGCGAAGTTGCAGCAAGCCGAGATCACACCTG 10589

58 aAlaProAlaAlaProPheArgProSerGlnGlyPheLeuSerArgL 75
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10588 TACTCCAGCC.....TGGCGCACAGACGACTGCTCTCAAAAAA 10545

75 euLysSerSerIleLysArgThrLysSerGlnProLysLeuAspArg 91
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10544 AAAAAA.....AAAGGAGAGTGGGAGGCC.....TGATCC 10507

92 SerSerPheArgGlnIleLeuProArgPheArgSerAlaAspHisAs 108
||:||||:||||:||||:||||:||||:||||:||||:||||:
10506 CTTCCTCTCTGAACTCTCTGCTGCC.....AG 10478

108 gAlaArgLeuMetGlnSerPheLysGluSerHisSerHisGluSer 125
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10477 GCGCCGCTGATGCAAGCTTTAAGGAGTCACTCTCATGATCTCTGC 10428

125 euserProSerSerAlaAlaGluAlaLeuGluLeuAsnLeuAspGlu 141
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10427 TGAGTCTCTAGCAGTGCAGCTGAGGCATTGGAGCTCACTTGGATG 10378

142 SerIleLysProValHisSerSerIleLysGlyGlnGluPheCys 158
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10377 TCCATTATCAAGCCAGTGCAGAGTCCATCTCTGGCCAGGAGTTCT 10328

158 eGlu..Val..... 160
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10327 TGAGTCTCTGCTGCTGGGCTGGGAGGCGCAAGACAGGGGTGATG 10278

160 ..... 160
10277 GAAAGTGGGGGAGAGAGGTCTTAGAGAAAGTGCACAGGTGGGACATC 10228

160 ..... 160
10227 AAGGAAACAGAAACTTCTGAGCTGGAGGAAAGGTAGCCAGGAGGAA 10178

160 ..... 160
10177 GAGAGGTAGCAGAGTCTCTCCCTCTGTAGCCCTTTCCCTCACTCCA 10128

161 .....ThrThrSerSerGlyThrLysCysPheAlaCysA 172
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10127 CACTCTTTCTAGTATCACTCTCATCAGGACAAATGCTTTGCTGTC 10078

172 rgSerAlaAlaGluArgAspLysTrpIleGluAsnLeuGlnArgAla 188
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10077 GGTCTGCGCCGAAAGACAGCAAAATGATGAGATCTGCAGCGGCGAGTA 10028

189 LysProAsn..... 191
||||:||||:||||:||||:||||:||||:||||:||||:||||:
10027 AAGCCACACAGGTATTGGGGAATAAGGGGACACAACTGTGCAGGCA 9978

191 ..... 191

```

9977 AAGGTTGCACCAACACGGGAAGCTGGGGCTAGGGAGGAAAGTGAGTTA 9928
 191 191
 9927 AAGGAGGAGAGGCTTGGGGAAGAGAGGATGAGGTACAGTGTATCTGGA 9878
 191 191
 9877 CAAGCAGGGGAGACCCCATTAATCTGAGTCCCCCATTTCTTTTCGCTT 9828
 191 191
 9827 TCTGACTGTACCTCCCTTACGATCTCTTCCCTGCCATAGAAGTCAAT 9778
 191 191
 9777 AGACTTACAGAGTTGATGGGGCTCTGGAAATTTCTTTAGTCTAGCTTCCCT 9728
 191 191
 9727 GCAGCAGGAGTCTTCACTAATACCCCGCAGGACACATCAATACACTT 9678
 191 191
 9677 AGCCAAAGTTTCTGTACCTTGGTTTCTCTCTAAGACAGAGGAACTGGTG 9628
 191 191
 9627 GCAATGGGTTGGATTAAATGATCTTAAGGTCCCTTTGGCACATAAATTC 9578
 191 191
 9577 TATGAGTCTGTCTTCCAAAACCTTCATGCTTCCAGTCTCTGTCTAGCTATT 9528
 191 191
 9527 CTCGTATCCAGTTTCTAGCAAGCTTACAGTCTAGTACACACTCCTCTG 9478
 191 191
 9477 GGGGGAAGCTTGTGTCTGTATGCTCTTAAGAAGCTCAACCAAGATGCT 9428
 191 191
 9427 GACATGATCTGACCTATGACAGGTACAAACACCATGTTCCCTTTCAACAC 9378
 191 191
 9377 ACTGCAATATCACCTGAGCCTTACCATTAGCTGTGCTCAATGTATCCC 9328
 191 191
 9327 CCCAGCCCTTACTGATCTTTTTTTTTTTTTTTTGGAGACAGTCTCAC 9278
 191 191
 9277 TCTGTACCCAGGCTGGAGTGAGTGGCGTGAATCTTGGCTCACTGCAACC 9228
 191 191
 9227 TCCGCTCTGGGTTTAAGCAATTCCTGTGCTCAGCCTTCCCCAGTAGCT 9178
 191 191
 9177 GGGATTACAGCATGCACCAACACACCCGACTAATTTTGTATTTTAGT 9128
 191 191
 9127 AGAGACAGGTTTCAACATGTTGGCAGGCTAGTCTCAAACCTCTGACCT 9078
 191 191

9077 CAGGTTGTCTGCTGCTCAGCCTCCCAAAGTCTAGGATTACAGGCATG 9028
 191 191
 9027 AGCCACCACATCTGGCCAGCTTTACCCATTCTATATATGCAATTTCTTTTT 8978
 191 191
 8977 CACAATTTCTGATAGTCTCTGACGAGCTTTCCAGTTCCCTTCAAATTTCTA 8928
 191 191
 8927 CATTAATATTTTGGCTTTTATTCAGCTTTTGAAATCTTTCCAAATACT 8878
 191 191
 8877 GTTTGTGTCTGTAGTATGTCTACGCTTCTACAGTAGCTTCTTGGGGC 8828
 191 191
 8827 TGCTATAACAACTGTACAACTGGGTCTCAGTCAAAACAAACAGAAATTA 8778
 191 191
 8777 TTCTCTCACAGTTCCAGAGGCTGGAAGTCCAAAGCAAGGTATCAGCAGG 8728
 191 191
 8727 GCCAGGCTCTTTTGATGCTGTAGCAGGGAATCCTTCCCTTGGCTCTTCT 8678
 191 191
 8677 AGCTTCCGAGGTTGCCAGAGTCTTGGCATTCCTTGGCTTTATAACTGC 8628
 191 191
 8627 ATCCGTCTAATTTCTGCTCCATCTTCATGTAGTGGCTTCTCTCTGTGT 8578
 191 191
 8577 GTCTGTATCTGTATCTGTCTGTCTCCTCAAAATCTCCCTCTCCATATAA 8528
 191 191
 8527 GACACAGTTTTTATAAGTGGTTAAGGTCCACTTAATTCAGTATGG 8478
 191 191
 8477 CCTTTTTTTTGAGATGGAATATCGCTCTTTTCCAGGCTTGGAGTG 8428
 191 191
 8427 CAATGGCATGAGTCTCAGTCTCAGTCAACCTCCGCTTCCAGGTTCAAGTG 8378
 191 191
 8377 ATTCTCTGCTCAGCCTCCTGAGTAGTGGGATTTACAGGCAGTGGCAC 8328
 191 191
 8327 TATGCCAGCTAATTTGTATTTTGTATTTTATTTAGACGGGTTTCACT 8278
 191 191
 8277 ATGTTGGCCAGGCTGGCTCGAACTCTGACCTCAGGCAATCCACCCACCT 8228
 191 191
 8227 CGGCTCCCAAAGTCTGGGATTACAGGAATGAGATPACCATGGCTGGCCT 8178
 191 191
 8177 CTCTCTAGTTGATTACATTTGCAGGGAATCTGTTTCCAAATAAGATCAC 8128

191 191
8127 ATTCACAGGTTTCAGGTAGACATGAGTTTGTGGGATACATATTCAACCC 8078
191 191
8077 AGTACACCGTCTCACCTTATGATTGCTGCTCTATCCACCTTCTTACTG 8028
191 191
8027 TCTCTCTCTTTTGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7978
192 191
192 Lys.. AspAsnSerArgValAspAsnValLeuLysL 204
7977 CACCCACCATGCCAGGACACACAGCCGCGGTAGACAATGTCTTAAGC 7928
204 eutrpillelleGluAlaArgGluLeuProProLysLysArgTyrTyrCys 220
7927 TGTGGATCATAGAGGCCCGGAGCTGCCGCCCAANGAGCGTACTACTGT 7878
221 GluLeuCysLeuAspMetLeuTyrAlaArgThrSerLysProAr 237
7877 GAGCTCTGCTGCATGACATGCTGTATGCACGCACCACTCCAAAGCCCG 7828
237 qSerAlaSerGlyAspThrValPheThrGlyGluHisPheGluPheAsn 254
7827 CTCTGCCTCTGGGGACACCGTCTTTCGGGGCGGCACCTTCGAGTTTACA 7778
254 snLeuProAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAsp 270
7777 ACCTGCCGGTCTGCCGTGCCCTGCCGCTGCATCTGTACCGTGACTCAGAC 7728
271 LysLysArgLysLysAspLysAlaGlyTyrValGlyLeuValThrValPr 287
7727 AAAAAGCGCAAGAAGACAGCAGCGTATGTCGGCCTGGTACTGTGCC 7678
287 ovalAlaThrLeuAlaGlyArgHisPheThrGluGlnTrpTyrProValt 304
7677 AGTGGCCACCTGGCTGGGCGGCACCTTCACAGACAGTGGTACCCCTGTAA 7628
304 hrLeuProThrGlySerGlyGlySerGlyGlyMetGlySerGlyGlyGly 320
7627 CCCTGCCAACAGCAGTGGGGATCTGGGGGCATGGTTCGGGAGGGGA 7578
321 GlyGlySerGlyGlySerGlyGlyLysGlyLysGlyGlyCysProAl 337
7577 GGGGGCTCGGGGGTGGCTCAGGGGGCAAGGCAAGGAGGTGGCCCGGC 7528
337 aValArgLeuLysAlaArgTyrGlnThrMetSerIleLeuProMetGluL 354
7527 TGTGGGCTCAAAAGCAGTTTACCAGACATGAGCATCTTGCCCATGGAGC 7478
354 eutYrLysGluPheAlaGluTyrValThrAsnHisTyrArgMetLeuCys 370
7477 TATATAAAGAGTTTCAGAGTATGTCACCAACCATATTCCGATGCTGTGT 7428
371 AlaValLeuGluProAlaLeuAsnValLysGlyLysGlyGluValAlaLase 387
7427 GCAGTCTTGAGCCCGCCCTGAATGTCAAAGGCAAGGAGGAGGTGGCAG 7378
387 rAlaLeuValHisIleLeuGlnSerThrGlyLysAla Lys 400
7377 TGCACATAGTTACATCTCCAGATACAGGCAAGGCAAGGCAAGGTGAGTGTG 7328
400 400
7327 TGCCCTCAGGAAAGGTGACTTGGGAATGGGCACCTTGCTTGGGGGTAGT 7278
400 400
7277 GAGGACAGGCAAAATTCACGAGATTGGGTGTGACAGAGCTTGACACITGG 7228

401 AspPheLeuSerAspMetAlaMetSerGluVal 411
7227 ATTTTCTGGCCCTCAGGACTTCTTTTCAGACATGGCCATGTCTGAGGTA 7178
412 AspArgPheMetGluArgGluHisLeuPheArgGluAsnThrLeuAl 428
7177 GACCGTTCATGGAACGGGAGCACCTCATATTCCGCGAGAACACAGCTTGC 7128
428 aThrLysAlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuL 445
7127 CACTAAAGCCCATAGAGAGTATATGAGACTGATTGGTCAGAAATACCTCA 7078
445 ysAspAla.Ile 448
7077 AGGATGCCATTGGTATGGCCACACACTCAGGCCCTCTCTCTCCAAACCTG 7028
448 448
7027 CCAGATGTCCACCCACAGACCCCAAGTCCACCTTCCACAGCTTGATACTT 6978
448 448
6977 CCTAACCCAGAGTCTTAGACTCCAGCTCCAAACACCTGATTCTGAAATT 6928
448 448
6927 TCCCAACCCCTGGCCACCCCTTCCCTGCCCTTGGAAAGTGTGACACAC 5878
449 GlyGluPheIleArgAlaLeuTyrGluL 458
6877 CCTCTGTGCCCCCACCACCCAGGAGAAATTCATCCGTCTGTATGAAT 6828
458 erGluGluAsnCysGluValAspProIleLysCysThrAlaSerSerLeu 474
6827 CTGAGAAACTGCGAGTAGACCTTATCAAGTGCACAGCATCCAGTTTG 6778
475 AlaGluHisGlnAlaAsnLeuArgMetCysGluLeuAlaLeuCysLys 491
6777 GCAGAGCACCAAGGCCAACCTGCGAATGTGCTGTGAGTTGGCCCTGTGCAA 6728
491 sValValAsnSerHis 496
6727 GGTGTCAACTCCCACTGGTGGAGACTGGGAACCTGGGCTGGGGGGCCAG 6678
496 496
6677 GGTGGGGGAATTTATGTTCATCTGTTCATCTATCTGTCCATCTCTCAAA 6628
496 496
6627 GAGGACTGAGCACCATTATTTATGGGCAAGCATTTCTTCTAGGCGCTATAGAG 6578
496 496
6577 CAACAGGTGAAGAGCCCTGGTCCCTGCCCTCAGAGGGCTCCACCAGA 6528
496 496
6527 ATGGGACAAATTAGAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6478
496 496
6477 GTGTAAGTGTGAGTAAGGTGCCCTCCCAACCTCTGTGTGACATAGGTCA 6428
496 496
6427 GAGAGAAGCAGAGCTTTGAGATAAGTGGGAAGAGGTGCCCTTGGGT 6378
496 496
6377 AGGCTTTGAAGACTGGTTAGGTCTGTATATGACATAGTTGGCAAGA 6328
496 496

6327 AAGACATTTTCAGAAGAAGGCTGTGAGAAAGCACATGTGTGATGTTGAAA 6278
496 496
6277 AGGCCAGAGATTTTCAGGGGACATTAAAGTAGGTTAGTACAAATACAT 6228
496 496
6227 CAGGTTAGTGGAGATGTGCCTCATATGAGGAGTGGGGGAGAGATGT 6178
496 496
6177 CTGGGCAGGAAGATTAGCTTTAGAACTGGAAGGCTCAAGAGTCTGAG 6128
496 496
6127 CTCATTGTTAGGCTTGGGATGCCATTAAAGGTGTCAAGAAATATTGTGAT 6078
496 496
6077 TTAGAAGATTAAATCCATAGCTGGGCACGGTGGCTCACACCTGTAATACC 6028
496 496
6027 AGCATTGAGGAGCTGAGCGGGCAGACCCTGAGGTCAGGAGTTTGA 5978
496 496
5977 GACCAGGCTGACCAACATGAGAAACCCCTGCTCTACTAAAAATACAAA 5928
496 496
5927 TTAGCCAGATGTGGTGGACATCCCTGTAAATCCAGCTATTGGGAGGCT 5878
496 496
5877 AAGCAGGAGAAATCATTGAATCTGGGAGTGGAGTGTAGTAGAGCGA 5828
496 496
5827 GATCACAGGATTGCATCTAGCTGGGCAACAAGAGTGAACCTCATCTC 5778
496 496
5777 AAAAAAAAAAATTAATCCATAAGTAGAGTGTATGTAGTAGAGTAGAGG 5728
496 496
5727 ATGTATGTTGGGGATGACTGGAAATAGAGCCAGCTGGGGGCTCGGAGCA 5678
496 496
5677 TGAAGTCAGATCCTGAATCAGAACAGTACAGAAAGTTAGGAAGGAGCTA 5628
496 496
5627 CTGGNAGACCCCTCTCGGAAAGGATCAGTAGGAATTGTCAGCTTATTGG 5578
496 496
5577 AAAGAGGGAGCATGTCTGTGGGAGTCAAGGATGACTCAAGAGGCCATG 5528
496 496
5527 AGGCTGGTGGTGGGACCGGTGGCCCTCATTCACAACCAAGGAAGTCAG 5478
496 496
5477 GAGGAGGAGCAAGTTGGAGGTGTGGGGCGATGGTGAAGTCTGCTTTACA 5428
496 496

5427 CCAGCCGAGTTTAAAGGTGTAGTGGGACATTGAAGTGAAGTGGGAGGTG 5378
496 496
5377 CGGGAGTAGGCTCTGAGCATTCAGGACTGGGGATCATGCCCTGGGGC 5328
496 496
5327 ACCTCCATCCCATTTCCCTGGAATCCAGAGAGTTGGGGGTCCGAGCT 5278
496 496
5277 CCCTGTACCTCAAGTGACCTCCATCTCTCTCCCATCTCTGTCTCTCCT 5228
496 496
5227 GGTGTCTGTTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5178
496 496
5177 CATCTCTCTCCAGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5128
496 496
5127 TCTCTCCTAATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5078
496 496
5077 GCCTCCCATCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5028
496 496
5027 GTTTCACACCTCACCTCTACACCCCTCACGATGTTCCCTGGAAG 4978
496 496
4977 CTGAGGCTCTCTGGGCTCAGTCCCGTCTCTCTCTCTCTCTCTCTCTCT 4928
497 CysValPheProArgGluLeuLysG 505
4927 CTCTGTCTCCCGACCTTCCCGGAGCTGTCTCCGAGGAGCTGAAG 4878
505 luAlaPheAlaSerTrpArgLeuArgCysAlaGluArgGlyArgGluAsp 521
|||||
4877 AGGTGTTTGTCTCTGGGCTGCGCTGGCAGAGCGAGCGGGAGGAC 4828
522 IleAlaAspArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCysPr 538
|||||
4827 ATCCGACACAGGCTTATCAGCGCTCCTCTCTCTCTCTCTCTCTCTCTCT 4778
538 oAlaIleMetSerProSerLeuPheGlyLeuMetGlnGluTyrProAspG 555
|||||
4777 AGCGATTATGTGCGCCAGTCTCTTTGGGCTTATGCAGAGTATCCCATG 4728
555 luGlnThrSerArgThrLeuIleAlaLysValIleGlnAsnLeu 571
|||||
4727 AGCAGCTCAGAAACCTCACCTCATTTGCCAAGTCTATCCAGAACCTG 4678
572 AlaAsnPheSerLys 576
4677 GCCAACTTTTCCAAAGTAGGGAAGTTCAGGAGTGGGCAGGCGAGGAGT 4628
576 576
4627 GGCAGGCGAGGAGTGGCAGGCTGGGGTTCGCAAGAAGGCTCTCTGA 4578
576 576
4577 GTCCCCCAGAGATCTGAGATGGGGAGGCTATGATACCTTGTGTGTGTG 4528
576 576
4527 TATGT 4478

```

577 .....PheThrSerLysGluAspPheLeuGlyP 586
4477 ATCTTCTGCAATCTTGCTAGTTCACCTCAAAGGAGGACTTCTCGGCT 4428
586 heMetAsnGluPheLeuGluGluTrpGlySerMetGlnGlnPheLeu 602
4427 TCATGAATAGTTCTGAGACTGGATGGGTTCCATGACGAGTTTGG 4378
603 TyrGluLeuSerAsnLeuAspThrLeuThrAsnSerSerPheGluG1 619
4377 TATGAGATCTCCAATCTGGACAGCTTAACCAACAGCAGTAGCTTTGAGGG 4328
619 yTyrIleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpG 636
4327 TTACATCGACTTGGCGGAGAGCTCTCCACACTGCATGCCCTACTCTGGG 4278
636 luValLeuProGlnLeuSerLys..Glu..... 644
4277 AGGTGCTGCCCGCAGCTCAGCAAGGTCAGCAGATCCOCTCTTTGCCCTATC 4228
644 ..... 644
4227 CCCAGATGGCTCCAGAGTTCTCTGGAGCCTGAGAACTACCTTTGAAGA 4178
644 ..... 644
4177 TTTTCTTTCTCCCTTCTTCTCGAGGTGTCACCACTACTATCCCAACTC 4128
644 ..... 644
4127 AGGCCCTCTCCACCTGCACCTCAGAGGCCCTTTAGAGCTGGGCACCTGA 4078
645 .....AlaLeuLeuLysLeuGly 650
4077 GCCCCAGGTAAACAGCTCACCTTCCAGGAAGCCCTCTGAAGCTGGGT 4028
651 ProLeuProArgLeuLeuSerAspIleSerThrAlaLeuArgAsnProAs 667
4027 CCACCTGCCCGGCTCTCTCAACGACATCAGCACAGCTCTCAGGAACCCCAA 3978
667 nIleGlnArgGlnProSerArgGlnSerGluArgAlaArgSerGlnProM 684
3977 CATCAAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3928
684 etValLeuArgGlyProSerAlaGluMetGlnGlyTyrMetMetArgAsp 700
3927 TGGTACTGCGGGGCCATCGCTGAGATGAGGAGGAGGAGGAGGAGGAGGAG 3878
701 LeuAsn..... 702
3877 CTCACAGGTGAGCACCCTGGGACAGCCAGGCTGTGCCCTAGGAGCCCT 3828
702 ..... 702
3827 TCTCTATTTCTAGATCTCTACTGGGCCCCACATGCACTCTCTAGGG 3778
702 ..... 702
3777 CTTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3728
702 ..... 702
3727 AGACAGGTACATGCTAGGTGAGGCTGAGAGCCCTTTACAGCCTGAGGAG 3678
702 ..... 702
3677 TGAGAGATTTGAGGCTCTAGGAATAGGGCTGAGGCTCCACCAACTCAGG 3628
702 ..... 702
3627 CTTAGTTGTAGCCTAGAGCATCCCTGCTGCAAGCTCTCAATTGCTGTGCC 3578
702 ..... 702
3577 CTCTGCTGCCCATGCTAGTCCCGAGGCTGAGGTTTCAGCCAGCATGTCAAT 3528
702 ..... 702
3527 GTACGCCATGTGTCAAAATGTTCAACATCTCAGTAATAGTAGTAGTAATA 3478
702 ..... 702
3477 AGCACTTCCCGCAGCCCGCCAGCACAAACCCACAGACCTCCCATGATCC 3428
702 ..... 702
3427 AGCACTTAGACAGTGCAGCAGAGCCTAGCAGGGCCTGCAGCCCTGCTC 3378
702 ..... 702
3377 CAGACTCCCGACCTCCATCTGATAGTGGTGGCCGTTGCTTTCATTGCTG 3328
702 ..... 702
3327 CTCAATTATTTGATGATGGCCCTGCTTTTCTCTGCCCCGTGCTTCTCTC 3278
702 ..... 702
3277 CATGACACATACCCATCCACCATTCGCCGCTCTCCTTTCATTGTTCCA 3228
703 .....S 703
3227 CATCTCTCTCTCTCTGCTGCTCGCCCTCTTTTCCATCTCTCTCCA 3178
703 erSerIleAspLeuGlnSerPheMetAlaArgGlyLeuAsn..... 716
3177 GCTCATCGACCTTCAAGTCTTCAAGTCTGAGGCCCTCAACAGGTGAGGG 3128
716 ..... 716
3127 GCTCTCCCTCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3078
716 ..... 716
3077 CTCACCTGGCCTTGGCCCTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 3028
716 ..... 716
3027 TCTCTCCATATGTGCCAGCCCTGCCCATCCCTCTCTCTCTCTCTCTCT 2978
716 ..... 716
2977 CATCTCCCTCTCTCTAGGCTTACCCCTTCCCGGAGGGCCCTGTCTT 2928
716 ..... 716
2927 TCCCTTTTACTACCTGTGCCCTCCCATCTCTCTCTCTCTCTCTCTCTCT 2878
716 ..... 716
2877 GCTGCCAGCGCTAGCTCTAGCCCTTCCCTCTCTCTCTCTCTCTCTCTCT 2828
716 ..... 716
2827 CAAGCCTGTGCCAGACACACAGCAAGGTTCAANTGTAGAGCCCTGACC 2778
717 .....Ser.SerMetAspMet 721
2777 TTACCTTCTGCTGTGTGCCCTTCCCTTCTCTGACAGCTCTATGGACATG 2728
722 AlaArgLeuProSerProThrLysGluLysProProProProProG1 738
2727 GCTGCCCTTCCCTTCCCAACCAAGAGAAAGCCACCCACCCCGCTGG 2678
738 yGlyGlyLysAspLeuPheTyrValSerArgProProLeuAlaArgSers 755

```

```
|||||
2677 TGGTGGTAAAGACCTGTTCTATGTAAGCGCTCCACCCCTGGCCGTTCTCT 2628
755 erProAlaTyrCysThrSerSerSerAspIleThrGluProGluGlnLys 771
|||||
2627 CACCAGCATCTGACAGGAGCTCGGACATCACAGAGCCAGAGCAGAAAG 2578
772 MetLeuSerValAsnLysSerValSerMetLeuAspLeuGlnGlyAspG1 788
2577 ATGCTGAGTGTCAACAAGAGTGTGTCATGCTGGACTTACAGGGGTGATGG 2528
788 yProGlyGlyArgLeuAsnSerSerValSerAsnLeuAlaAlaValG 805
|||||
2527 GCCTGGTGGCGGCTCAACAGCAGCAGTGTTCGAACCTGGGGGCCGTAG 2478
805 lAspLeuLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeu 821
2477 GGGACCTGCTGCATCAAGCCAGGCTCGCTGACAGCAGCCTTGGGGCTA 2428
822 ArgProAlaProAlaGlyArgLeuSerGlnGlySerGlySerSerIleTh 838
|||||
2427 CGGCTGGCGCTCGCGGAGCCCTCTCCAGGGGAGTGGCTATCATCATCAC 2378
838 rAlaAlaGlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValP 855
2377 GCGCGCTGGCATGGCCCTCAGCCAGATGGGTGTACACACAGCAGCGTGTCC 2328
855 roAlaGlnGlnLeuArgIleProLeuSerPheGlnAsnProLeuPheHis 871
|||||
2327 CTGCCCAGCAACTGGGAATCCCTCTCTCTCCAGAACCTCTCTTCCAC 2278
872 MetAlaAlaAspGlyProGlyProProAlaGlyHisGlySerSerG1 888
|||||
2277 ATGGCTGCTATGGGCGAGCTCCCGGAGCCCTTCCCGGCGCATGGAGGGGCGGTGG 2228
888 yHisGlyProProSerSerHisHisHisHisHisHisHisHisHis 905
2227 CCATGGGCCACCTTCTCTCCATCACCACCACCACCACCACCACCACCACC 2178
905 rGlyGlyGluProProGlyAspThrPheAlaProPheHisGlyTyrSer 921
2177 GAGGTGGAGAGCCCTTGGGGAGCACCTTGGCCCATTCATGGCTATAGC 2128
922 LysSerGluAspLeuSerThrGlyValProLysProProAlaAlaSerI1 938
2127 AAGAGTGAGGACCTCTCTCGGGGTCCCGCAAGCCCTCTGCTGCCATC 2078
938 eLeuHisSerHisSerTyrSerAspGluPheGlyProSerGlyThrAspP 955
2077 CTTTCATAGCCACAGCTACAGTGTATGAGTTTGGACCTCTGGCACTGACT 2028
955 heThrArgArgGlnLeuSerLeuGlnAspAsnLeuGlnHisMetLeuSer 971
2027 TCACCCGTGGCAGCTTTCATCTCAGAGAACCTTCGACGACATGCTGTCC 1978
972 ProProGlnIleThrIleGlyProGlnArgProAlaProSerGlyProG1 988
1977 CCTCCCAAGATCACCATTGTCCTCCAGAGCCAGCCCTCAGGGCCCTGG 1928
988 yGlyGlySerGlyGlySerGlyGlyGlyGlyGlyGlyGlnProProp 1005
1927 AGGTGGGAGCGGTGGGGGAGCGGTGGGGTGGGGGGCCAGCGCCCTC 1878
1005 roLeuGlnArgGlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLys 1021
1877 CATTCGACAGGGGCAAGTCTCAGCAGTTGACAGTCAGCGCAGCCCGAGAA 1828
1022 ProArgProSerSerGlyAsnLeuLeuGlnSerProGluProSerTyrG1 1038
1827 CCCCCGCCATCCAGCGGGAATCTATTGCACTGCCAGAGCCCAAGTATGG 1778
1038 yProAlaArgProArgGlnGlnSerLeuSerLysGluGlySerIleGlyG 1055
|||||
```

```
1777 CCGCGCCCTCCACGGCAACAGAGCGCTCAGCAAGAGGAGGCGAGCATTTGGG 1728
1055 lySerGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1071
|||||
1727 GCACGGGGGCGACGGGTGGCGGAGGGGTGGGGGCTGAGGCCCTCCATC 1678
1072 ThrLys 1073
|||||
1677 ACCAAGCAGGTAGGTGAAGSCAGAGGAGGAAGGCGGCTGGGTCAACAACAG 1628
1073 1073
1627 GAGGGAAGAGAGATGGGGGTGGGGTTGAAACAGAGTCTGTGCGCTGA 1578
1073 1073
1577 AGTTACAATCTTCTGCTGCTCTTGGCCATTAAACAAATGTGTATAGAG 1528
1073 1073
1527 GGCCTGCTATGTCCTATGCTATGCCAGGCACTAAGGATACGGCACTGA 1478
1073 1073
1477 ACCCTCTTAGCACTCTCATTTCCAGAGGGGATTAAATCCATAAGTAGAATGG 1428
1073 1073
1427 GGGTGACTGGAATAGGCGGAGCTGGGGGCTTGGAGGCGATGAAGATCAGAT 1378
1073 1073
1377 CCTGAATCAGAACACAGGAGAAAATTTCTCTGGCATTCAACTCACATCTCT 1328
1073 1073
1327 GGCATTTCAGAGGTGATAAAGGCTTACAGCAGGCAATAGGACTTGGGACTGT 1278
1073 1073
1277 CGTCTCTTTGGCTGTGCTGTTGCCCTCTAAATGTACCTGCTTTTTCCTCC 1228
1074 1074
1227 ACCTTTTCTTCTGTTGCCCTCTACTGTGCTTGTGCCAGCATTTCTCA 1178
1077 nThrProSerThrLeuAsnProThrMetProAlaSerGluArgThrValA 1094
|||||
1177 GACACCATCCACATTGAACCCCAANTGCCAGCTCTGAGCGGACAGTGG 1128
1094 laTrpValSerAsnMetProHisLeuSerAlaAspIleGluSerAlaHis 1110
|||||
1127 CTTGGGTCTCCAACTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1078
1111 lIleGluArgGluGluTyrLysLeuLysGluTyrSerLysSerMetAspG1 1127
|||||
1077 ATCGAGCGGGAAGAGTACAAGCTCAAGGAGTACTCAAAATTCATGATGATGA 1028
1127 userArgLeuAspArgValLys 1134
|||||
1027 GAGCGCGCTGGATAGGATACGA 1006
```

seq_name: gb_htg20:AL161906

seq_documentation_block:

LOCUS AL161906 169323 bp DNA HTG 18-AUG-2000
DEFINITION Homo sapiens chromosome 6 clone RP11-567N9, *** SEQUENCING IN
PROGRESS ***, 24 unordered pieces.

ACCESSION AL161906
VERSION AL161906.5 GI:9863619
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 169323)

REFERENCE AUTHORS TITLE JOURNAL

Submitted (17-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:9212890.

COMMENT

----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BA567N9

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: plasmid; L08752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 155424 bases at least Q40

Consensus quality: 160003 bases at least Q30

Consensus quality: 163252 bases at least Q20

Insert size: 167023; sum-of-contigs

Insert size: 186175; 5.7% error; agarose-fp

Quality coverage: 3.59x in Q20 bases; sum-of-contigs Quality

coverage: 3.22x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 5456: contig of 5456 bp in length
* 5457 5556: gap of 100 bp
* 5557 8767: contig of 3211 bp in length
* 8768 8867: gap of 100 bp
* 8868 17168: contig of 8301 bp in length
* 17169 17268: gap of 100 bp
* 17269 26591: contig of 9323 bp in length
* 26592 26691: gap of 100 bp
* 26692 31188: contig of 4497 bp in length
* 31189 31288: gap of 100 bp
* 31289 37579: contig of 6291 bp in length
* 37580 37679: gap of 100 bp
* 37680 45350: contig of 7671 bp in length
* 45351 45450: gap of 100 bp
* 45451 48354: contig of 2904 bp in length
* 48355 48454: gap of 100 bp
* 48455 63401: contig of 14947 bp in length
* 63402 63501: gap of 100 bp
* 63502 67724: contig of 4223 bp in length
* 67725 67824: gap of 100 bp
* 67825 70550: contig of 2726 bp in length
* 70551 70650: gap of 100 bp
* 70651 80664: contig of 10014 bp in length
* 80665 80764: gap of 100 bp
* 80765 83307: contig of 2543 bp in length
* 83308 83407: gap of 100 bp
* 83408 88383: contig of 4976 bp in length
* 88384 88483: gap of 100 bp
* 88484 93697: contig of 5214 bp in length
* 93698 93797: gap of 100 bp
* 93798 99354: contig of 5557 bp in length
* 99355 99454: gap of 100 bp
* 99455 111055: contig of 11601 bp in length
* 111056 111156: gap of 100 bp
* 111156 118645: contig of 7490 bp in length
* 118646 118745: gap of 100 bp
* 118746 122421: contig of 3676 bp in length

* 122422 122521: gap of 100 bp
* 122522 132468: contig of 9947 bp in length
* 132469 132568: gap of 100 bp
* 132569 137959: contig of 5391 bp in length
* 137960 138059: gap of 100 bp
* 138060 151691: contig of 13632 bp in length
* 151692 151791: gap of 100 bp
* 151792 158253: contig of 6462 bp in length
* 158254 158353: gap of 100 bp
* 158354 169323: contig of 10970 bp in length.

FEATURES

Source

1..169323

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="6"

/clone="RP11-567N9"

/clone_lib="RPC1-11.2"

1..5456

/note="assembly_fragment:00884"

fragment_chain:1"

5557..8767

/note="assembly_fragment:00832"

fragment_chain:1"

8868..17168

/note="assembly_fragment:01060"

fragment_chain:1"

17269..26591

/note="assembly_fragment:01172"

fragment_chain:1"

26692..31188

/note="assembly_fragment:00279"

fragment_chain:1"

31289..37579

/note="assembly_fragment:01205"

fragment_chain:1"

37680..45350

/note="assembly_fragment:01334"

fragment_chain:2"

45451..48354

/note="assembly_fragment:01263"

fragment_chain:2"

48455..53401

/note="assembly_fragment:00990"

fragment_chain:2"

63502..67724

/note="assembly_fragment:00045"

fragment_chain:2"

67825..70550

/note="assembly_fragment:00965"

fragment_chain:2"

70651..80664

/note="assembly_fragment:00637"

fragment_chain:3"

80765..83307

/note="assembly_fragment:01291"

fragment_chain:3"

83408..88383

/note="assembly_fragment:01317"

fragment_chain:3"

88484..93697

/note="assembly_fragment:01171"

fragment_chain:4"

93798..99354

/note="assembly_fragment:00288"

fragment_chain:4"

99455..111055

/note="assembly_fragment:00286"

fragment_chain:4"

111156..118645

/note="assembly_fragment:00303"

fragment_chain:5"

118746..122421

/note="assembly_fragment:01184"

```

misc_feature      fragment_chain:5'
122522..132468
/note="assembly_fragment:00501
fragment_chain:6'
132569..137959
/note="assembly_fragment:01244
fragment_chain:6'
138060..151691
/note="assembly_fragment:00032"
151792..158253
/note="assembly_fragment:00779"
158354..169323
/note="assembly_fragment:01270"
169424..179899
BASE COUNT 41674 a 41332 c 41269 g 42673 t 2375 others
ORIGIN

alignment_scores:
  Quality: 3987.00      Length: 2457
  Ratio: 4.237         Gaps: 12
Percent Similarity: 38.299 Percent Identity: 37.810

alignment_block:
US-09-294-298-4 x AL161906 ..
Align seg 1/1 to: AL161906 from: 1 to: 169323

190  ProAsnLysAspAsnSerArgArgValAspAsnValLeuLysLeuTyrPIL 206
||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
77990 CCATGCCAGGACACACGCCCGGGGTAGACATGCTGCTAAAGCTGTGGAT 78039

206  eileGluAlaArgGluLeuProProLysLysArgTyrTyrCysGluLeuC 223
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78040 CATAGAGCCCGGAGGTGCGCCCGCCCAAGAGCGGTACTGTGAGCTCT 78089

223  ysLeuAspAspMetLeuTyrAlaArgThrThrSerLysProArgSerAla 239
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78090 GCCTGGATGATGCTGTATGACGACGACACCTCCAAAGCCCGCTCTGCC 78139

240  SerGlyAspThrValPheThrPGLyGluHisPheGluPheAsnAsnLeuPr 256
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78140 TCTGGGACACCGCTCTCTGGGGCGGACACTTCGAGATTAAACCTGCC 78189

256  oAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysA 273
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78190 GGCTGTCCGTGCCCTGCGGCTGCATCTGTACCGTGACTCAGACAAAAGC 78239

273  rGLyLysAspLysAlaGlyTyrValGlyLeuValThrValProValAla 289
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78240 GCAAGAGGACAGGCGGCTATGTCTGGCTGTGCTGCTGCTGCTGCTGCC 78289

290  ThrLeuAlaGlyArgHisPheThrGluGlnTyrTyrProValThrLeuPr 306
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78290 ACCCTGGCTGGCGGCCACTTCACAGACAGTGTACCTGTAAACCTGCC 78339

306  oThrGlySerGlyGlySerGlyGlyMetGlySerGlyGlyGlyGlyGly 323
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78340 AACAGGACAGTGGGGATCTGGGGGATGGTTCCGGAGGGGGGGGGGCT 78389

323  erGlyGlyGlySerGlyGlyLysGlyGlyGlyGlyGlyGlyGlyGly 339
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78390 CGGGGGTGGCTCAGGGGGGCAAGGGAAGAGGTGGCCCGCTGTGGCG 78439

340  LeuLysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeuTyrLy 356
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78440 CTGAAGCAGCTTACACAGCAATGACCATCTTGCCCATGGAGCTATATA 78489

356  sGluPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaValL 373
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78490 AGAGTTTGCAGAGTATGTCACCAACCATTTATCGGATGCTGTGCGAGTCT 78539

373  euGluProAlaLeuAsnValLysGlyLysGluValAlaSerAlaLeu 389
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

78540 TGGAGCCCGCCCTGAATGTCAAAGCAAGAGGAGGTTGCCAGTGCACATA 78589
390  ValHisIleLeuGlnSerThrGlyLysAla Lys..... 400
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78590 GTTCACATCTCGAGAGTACAGCAAGGCCAAGGTGAGTGTGTGCCCTC 78639
400 ..... 400
78640 AGGAAAGGTGACTTGGGAATGGGACCTTGTCTGGGGGTTAGTGAGGACA 78689
400 ..... 400
78690 GGGCAAAATTCACAGATTGGGTGTGTCAGAGGCTCACACATTGGATTTC 78739
401 ..... AspPheLeuSerAspMetAlaMetSerGluValAspArgP 414
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78740 TGGGCCCTCAGGACTTCCCTTTTCAGACATGGCCATGCTCAGGTAGACCGT 78789
414  heMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAlaThrLys 430
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78790 TCATGGAACGGGAGCACCTCATATTCGCGAGACACACGCTTGCCACTAAA 78839
431  AlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuLysAspAl 447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78840 GCCATAGAGAGTATATGAGACTGTGTGTGTCAGAAATACCTCAAGGATGC 78889
447  a.Ile..... 448
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
78890 CATTTGGTATGGCCACACACTCAGGCCCTCTCTTCTCCAAACCTGCCAGATG 78939
448 ..... 448
78940 TCCACCCACACCCCAAGTCCACCCCTTCCACAGCTTGATACTTCTTAACC 78989
448 ..... 448
78990 CAGAGTCTTAGGACTCCAGCTCCACACCTGATTTCTGAAATTTCCCCAA 79039
448 ..... 448
79040 CCTTGGCCACCCCTTCCCTTGGCCCTTGGAAAGTGTGACCACACCTCTTTG 79089
449 ..... GlyGluPheIleArgAlaLeuTyrGluSerGluG 460
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
79090 TGCCCCCACCACCCAGGAGAAATTCATCCGTGCTGTATGATCTGAGGA 79139
460  uAsnCysGluValAspProIleLysCysThrAlaSerSerLeuAlaGluH 477
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
79140 AAATGCGAGTAGACCTTATCAAGTGCACAGCATCCAGTTTGGCAGAGC 79189
477  isGlnAlaAsnLeuArgMetCysGluLeuAlaLeuCysLysValVal 493
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
79190 ACCAGGCAACCTGCGCAATGTGCTGTGAGTTGGCCCTGTGCAAGGTGTC 79239
494  AsnSerHis..... 496
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
79240 AACTCCCACTGGTGAGACTGGGAACCTTGGGCTGGGGGCCCGGCTCGGG 79289
496 ..... 496
79290 GGAATTATGTTCATCTGTTTCATCTATCTGTCATCTGTCATCTCAAGAGGACT 79339
496 ..... 496
79340 GAGCACCATTTTGGGCAAGCATTTGTTCTAGGCGCTATAGAGCAACAG 79389
496 ..... 496
79390 GTGAAAGAGGCGCTGGTCCCTGCCCCCTCAGAGGCGCTCCACCAGATGGGA 79439
496 ..... 496
79440 CAAATTAGAGAAAAAAGGCAAGGCAAGGCAAGGCAATATGTTGTGTAAAG 79489

```

```

496 ..... 496
79490 TGCTGAGTAAGGTCCCCCAACCTCTGTGACATAAGGTCAGAGAGAA 79539
496 ..... 496
79540 GGCAGAGCTTTGAGATAAGTGGGGAAGAGGTGCCCCCTTGGGTAGGCTTT 79589
496 ..... 496
79590 GAAGACTGGTTTAGTCTGATATATGGACATAGTTGGCAAGAAAGACAT 79639
496 ..... 496
79640 TTCAGAGAAGGCTGTGAGNAAGGCACATCTGTGATGGTGAAAGGCCCA 79689
496 ..... 496
79690 GGAGTTTTTCAGGGGACATTAAAGTAGGTTAGTAGCAATTACATCAGGTTT 79739
496 ..... 496
79740 AGTGGAGCATGTCCTCATATAATGGGAGTGGCGGAGAGATCTCTGGCA 79789
496 ..... 496
79790 GGAAGATTAGCTTTAGAAACTGGAAGGCCTCAAGGAGTCTGAGGTCAATTG 79839
496 ..... 496
79840 TTAGGCCTTGGGATGCCATTAAAGGTGTCAGAAATATTGTGATTTAGAAG 79889
496 ..... 496
79890 ATTAACTCATTAGCTGGGCAAGCTGTGCTCACACCTGTAATACCAAGCACTT 79939
496 ..... 496
79940 TGGGAGCTGAGGGGCGAGACCACCTGAGGTGAGGATTTGAGACCAGG 79989
496 ..... 496
79990 CTGACCAACATGGAGAAACCTCTCTCTACTAATAACAAATTAGCCA 80039
496 ..... 496
80040 GATGTGTGGACATGCTCTGTAATCCAGCTATTCGGGAGGCTAAGGCAG 80089
496 ..... 496
80090 GAGAAATCATTGAATCTGGGAGGTGGAGGTTGTAGTGACCGAGATCACA 80139
496 ..... 496
80140 CCATTGCACCTAGCTGGGCAACAAGAGTGAATCCGCTCBAAAAAA 80189
496 ..... 496
80190 AAAAAATTAATCCATAAGTAGAGTGTATGTAGTAGAGTAGAGGATGTATG 80239
496 ..... 496
80240 TTGGGGATGACTGGAATAGACCGACCTGGGGCTCGGAGGCATGAAGT 80289
496 ..... 496
80290 CAGATCCTGTAATCAGAACAGTACAGAAAGTTAGGAAGGAGCTACTTGAAG 80339
496 ..... 496
80340 GACCCCTCTCGGAAAGGATCAGTAGGAATGTGACGTTATTGGAAGGAG 80389

```

```

496 ..... 496
80390 GGAGCATCTCTGTGGGAGTCAAGGATGACTCAAGAGGCCATGAGGCTGG 80439
496 ..... 496
80440 TGGTTGGAGACCGGTGGCCTCATTTGACAAACAGGAAAGTCAGGAGGAGG 80489
496 ..... 496
80490 AGCCAGTTGGAGGTGTGGGGCGATGTGAGTCTGCTTTTACACACCGCG 80539
496 ..... 496
80540 AGTTTAAAGGTCTCAGTGGGACATTTGAAGTGAACCTGGGAGTGGGGGAG 80589
496 ..... 496
80590 TGAGCTCTGAGCCATTCCAGGAGTGGGATCATGCTGGGGCACTCCA 80639
496 ..... 496
80640 TCCCATTTCCCTGGATCCAGAAAGANNNNNNNNNNNNNNNNNNNNNNNN 80689
496 ..... 496
80690 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 80739
496 ..... 496
80740 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 80789
496 ..... 496
80790 GGTCGGAGCTCCCTGTACCTCAAGTGACCTCCATCTCTCTCCCATCTCT 80839
496 ..... 496
80840 GTCTCTCCCTGGTGTCTGTTTTTCTCTCTCTCTCTCTCTCTCTCTCTCC 80889
496 ..... 496
80890 CACACCCCTCCATCTCTCCAGGTGTCTCTCCCTCACCCTCTCTCTCTCC 80939
496 ..... 496
80940 CCTCATTTCTCTCTCTCCCTAATCTGTCTGTTCCTCTGCATGGCCCCCT 80989
496 ..... 496
80990 TCTTCAAGCAGCTCCCATCTGTCTCTGGGTCTCTCTCTCTCTCTCTCTCT 81039
496 ..... 496
81040 TCTCACCCCTGTTTCCACACCCCTCACCTCTCTACACCCCTCAGCATGT 81089
496 ..... 496
81090 TCCCTGGAAGCTGAGGCTCTCTGGGGCTCAGTCCCGTCTCTCTCTCTCT 81139
497 ..... 501
81140 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 81189
502 GluLeuLysGluValPheAlaSerTrpArgLeuArgCysAlaGluArgI 518
81190 GAGCTGAAGAGGAGTGTGTCTTTCATGGCGCTGCTGCGCAGAGCGAGG 81239
518 YargGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuArgp 535
81240 CCGGAGGACATCGACAGAGGCTTATCAGCGCTCACCTCTCTCTCTCTCT 81289
535 heLeuCysProAlaIleMetSerProSerLeuPheGlyLeuMetGlnGlu 551

```

|||||
81290 TCCTCTGCCAGCGATTATGTGCCCGAGTCTTTGGGCTTATGCAGGAG 81339
552 TyrProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysValIle 568
|||||
81340 TACCAGATGACGACACCTCACGAACCTCCATCCCTCATTTGCCAAGGTCAT 81389
568 eGlnAsnLeuAlaAsnPheSerLys. 576
|||||
81390 CCAGAACCTGGCCAACCTTTTCCAAGTGAGGGAAGCTTCAGAGTGGCGAG 81439
576 576
81440 GGCAGGAGTGGCAGGCGAGGAGTGGCAGGCGTGGGGTGGCAAGAAG 81489
576 576
81490 GGTCTCCTGAGTCCCGAGAGATCCTGAGATGGGAGGCTATGATACCTTG 81539
576 576
81540 TGTGTGTGTATGTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 81589
577 PheThrSerLysGluasp 582
|||||
81590 TGTGACCTTTATCTCTGCAATCTTTGGCTAGGCTTTACCTCAAGGAGGAC 81639
583 PheLeuGlyPheMetAsnGluPheLeuGluLeuGluTrpGlySerMetG1 599
|||||
81640 TTTCTGGGCTTCATGAATGAGTTCTCGAGCTGGAATGGGTTCCATGCA 81689
599 nGlnPheLeuTyrgluIleSerAsnLeuAspThrLeuThrAsnSerSerS 616
|||||
81690 GCAGTTTTTGTATGAGATCTCCAATCTGGACACGCTCAACACAGCAGTA 81739
616 erPheGluGlyTyrrileAspLeuGlyArgGluLeuSerThrLeuHisAla 632
|||||
81740 GCITTTAGGGTTACATCGACTTTGGGCGGAGAGCTCTCCACACTGCATGCC 81789
633 LeuLeuTrpGluValLeuProGlnLeuSerLys...Glu. 644
81790 CTACTCTGGAGGTGTGCCCCAGCTCAACAAGGTCAAGAGTATCCCTCT 81839
644 644
81840 TTGCCCTATCCCGAGATGGCTCCAGAGTTCTGGAGGCTGAGAACTAC 81889
644 644
81890 CTTTGAAGATTTTTCCTCCCTTGTTCGAGGTGTACACCCTACT 81939
644 644
81940 ATCCCAACTCAGCCCCCTCCACCTGCACCCCTCAGAGGCCCTTTAGAGC 81989
645 AlaLeuLe 647
81990 TGGGCACTAGCCCCCAGGTAACAGGCTCACCCCTCCAGGAGCCCTCT 82039
647 uLysLeuGlyProLeuProArgLeuLeuSerAspIleSerThrAlaLeuA 664
82040 GAAGCTGGTCCACTGCGCGGCTCTCAACGACATCAGCAGCTCTGA 82089
664 rAsnProAsnIleGlnArgGlnProSerArgGlnSerGluArgAlaArg 680
82090 GGAACCCCAACATCCAAGGACGCCAAGCGCCAGAGTGAAGGCCGCCGG 82139
681 SerGlnProMetValLeuArgGlyProSerAlaGluMetGlnGlyTyrMe 697
82140 CCTCAGCTGTGGTACTGCGGGGCCATCGGCTGAGATCAGGGCTACAT 82189
697 tMetArgAspLeuasn. 702
|||||

82190 GATCGGGACCTCAACAGGTGACACCCCTGGGACAGCCAGGCTGTGCC 82239
702 702
82240 TAGGAGCCCTTCTCTATTCTAGATACTCTCAGTGGGCCACACATGCAT 82289
702 702
82290 CTCCTAGGCTTGAAGAAGGAGGAAAAAGCACCAGTTCTCAGGGGA 82339
702 702
82340 GACGATAAGGAGACAGGTACAGTACAGTGTAGGCTGAGAGCCCTTTACAG 82389
702 702
82390 CCTGAGGAGTGAGAGATTGGAGCTCTAGGAATAGGGCTGAGGCTCCAC 82439
702 702
82440 CAACCTACGGCTTAGTTGTAAGCTAGAGCATCCCTGCTGCAAGCTCTGA 82489
702 702
82490 TTTGCTGTCCCTGTGCTGCCATGTAGTCCCGGCTGAGGTTTCAGCC 82539
702 702
82540 AGCATGTCATGTCAGCCATGTGTAAAAATGTTCAAAACATCTCAGTAATAG 82589
702 702
82590 CTAGTGAATAAGCACACTTCCCGGAGCCCGGACCAACCCACAGACCTC 82639
702 702
82640 CCCATGATCCAGCACTTAGAGCAGTGACAGCAGAGCCTAGCAGGGCTG 82689
702 702
82690 CAGCCTGCTCCAGAGTCCAGCCCTCCATTCGTAGTGGTGGCCGCTGT 82739
702 702
82740 TCATTTGCTGCTCATTTATTTGATGATGGCCCTGCTTTTCTCTCTGCCG 82789
702 702
82790 TGTCTTCTCCATGACACCATACCCATCCACCATTCGCCCTCTCCTTT 82839
702 702
82840 CATTTGTCCACATCTCTCTCTCTGCTGTGCTGCTGCCCTCTTTTCCA 82889
703 SerSerIleAspLeuGlnSerPheMetAlaArgGlyLeuAs 716
82890 TCTCTCCAGCTCCATCGACCTTCAGTCTTCATGGCTCGAGCCCTCAA 82939
716 n. 716
82940 CAGGTGAGGGCTCTCCCTCCCGCCCTCTCTCTCTCTCTCTCTCTCTGT 82989
716 716
82990 CCCTCTCCACTCCACTGGGCTTGGCCCTTACTCTCTCTCTCTCTCTCTC 83039
716 716
83040 ATGGACCTCATCTCTCTCCATATGTGCCAGCCCTGCCCATCCCTCTCTC 83089
716 716
83090 TTGCTGCCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 83139

716 716
83140 CCCTGTCTTTTCCCTTTACTTACCTGTGCTCCCTCCCTTCCCTCCCTGCTG 83189
716 716
83190 CTCTTTTATAGGCGTGTCAACCGTTAGCTCTCAGCCCTCCCTCTGGTCCCC 83239
716 716
83240 TTTTACCAGGCGCTTGTGCCACCAACAAGGTCAATTGTAGGGA 83289
716 716
83290 GCCCTTACCTAACCCTTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 83339
716 716
83340 NNN 83389
716 716
83390 NNN 83439
716 716
83440 ACTCACCTGTCCCTCCCATCTCCCTGCCCTGCCCTCTTCAGGCGTCCCA 83489
716 716
83490 CGCTAGCTCTCAGCCCTTCCCTCTGGGTCCCACTTTTACCCCAAGGCC 83539
716 716
83540 TGTGCCAGACCACAGAGGTTCAATTGCTAGGAGCCCTGACCTTACCTT 83589
717 Ser...SerMetAspMetAlaArgL 724
83590 CTGCTGTGTGCGCCCTTCCCTTTCAGACGCTCTATGGACATGGCTGCC 83639
724 euProSerProThrLysGluLysProProProProProProGlyGly 740
83640 TCCCTTCCCAACCAAGGAAAGCCACCCACCCACCCCTGGTGGT 83689
741 LysAspLeuPheTyxValSerArgProProLeuAlaArgSerProAl 757
83690 AAGACCTGTCTATGTAGCCGTCCACCCCTTGGCCCGTTCCCTCACCAGC 83739
757 atyrcysThrSerSerAspIleThrGluProGluGlnLysMetLeuS 774
83740 ATACTGCAGCAGCAGCTCGACATCACAGAGCCAGAGCAGAAGATGCTGA 83789
774 erValasnLysSerValSerMetLeuAspLeuGlnGlyAspGlyProGly 790
83790 GTGTCAACAAGAGTGTTCATGTGGACTTACAGGCTGATGGCCCTGGT 83839
791 GlyArgLeuAsnSerSerValSerAsnLeuAlaValGlyAspLe 807
83840 GGCCGCTCAACAGCAGCAGTGTTCGAACCTGGCGCCGTAGGGGACCT 83889
807 uLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProA 824
83890 GCTGCACCTAAGCCAGGCGTCTCCAGAGGAGTGGCTCATCATCAGCGCGCT 83939
824 laProAlaGlyArgLeuSerGlnGlySerGlySerIleThrAlaAla 840
83940 CGCTGCGGAGCGCTCTCCAGAGGAGTGGCTCATCATCAGCGCGCT 83989
841 GlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaG 857
83990 GGCATGCGCTCAGCCAGATGGGTGTCCACACAGAGCGTGTCCCTGCCCA 84039

857 nglnLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaA 874
84040 GCAACTGCGAATCCCTCTCCCTCCAGAACCTCTCTTCCACATGGCTG 84089
874 laAspGlyProGlyProProAlaGlyHisGlyGlySerSerGlyHisGly 890
84090 CTGATGGCCAGGTCCCCAGCGGCCCATGGAGGGCGGTGGCCATGGC 84139
891 ProProSerSerHisHisHisHisHisHisHisHisHisHisArgGlyG 907
84140 CCACCTCTCTCCCATCACACACACACACACATCACACACACAGGTGG 84189
907 yGluProProGlyAspThrPheAlaProPheHisGlyTySerLysSerG 924
84190 AGAGCCCTCTGGGACACCTTGGCCCATTCCTCCCTCCATCTTCAT 84239
924 luAspLeuSerThrGlyValProLysProProAlaAlaSerIleLeuHis 940
84240 AGGACCTCTCTCGGGGTCCCAAGCCCTGCTGCCCTCCATCTTCAT 84289
941 SerHisSerTySerAspGluPheGlyProSerGlyThrAspPheThrAr 957
84290 AGCCACAGCTACAGTATGATGTTGGACCTCTGGCAGCTGACTTCACCG 84339
957 gArgGlnLeuSerLeuGlnAspAsnLeuGlnHisMetLeuSerProProG 974
84340 TCGCAGCTTCTACTCCAGGACAACCTGCAGCACATGCTGTCCCTCCCC 84389
974 inIleThrIleGlyProGlnArgProAlaProSerGlyProGlyGly 990
84390 AGATCACCATTTGTCAGAGCCAGCCCTCAGGGCTTGAGGTGG 84439
991 SerGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlnProProProLeuG 1007
84440 AGCGTGGGGGACGCGTGGGGTGGCGGGGCCAGCGCTCCATTCGA 84489
1007 nArgGlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLysProArgP 1024
84490 GAGGGCAAGTCTCAGCAGTGTACAGTCAGCGAGCCAGCCAGAACCCCGC 84539
1024 roSerSerGlyAsnLeuLeuGlnSerProGluProSerTySerGlyProAla 1040
84540 CATCAGCGGGAATCTATTGACGTCCAGAGCCAAAGTTATGGCCCGCC 84589
1041 ArgProArgGlnGlnSerLeuSerLysGluGlySerIleGlyGlySerCl 1057
84590 CGTCCAGCAACAGAGCCCTCAGAGGAGGCGAGCATTTGGGGGAGCG 84639
1057 yGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyLeuLysProSerIleThrLys 1073
84640 GGGCAGCGGTGGCGAGGGGTGGGGGCTGAAGCCCTCCATCACCAGC 84689
1073 1073
84690 AGGTAGGTGAAGCAGGAGGAGCGGGGTGGGTCAACAAGGAGGAA 84739
1073 1073
84740 GAAGGAGATGGGGGTGGGTGTGAACAGAGTCTGTGGCTGAAGTTACA 84789
1073 1073
84790 ATCTTCTTGCTCTCTTTTGGCCATTAAACAAATGTGTATAGAGGCCCTGC 84839
1073 1073
84840 TATGTGCCATGTGCTATGCCAGCAGCTAAGATACGCACTGAACCTCT 84889
1073 1073
84890 TAGCACTCTCATCCAGAGGGGATTAATCCATAAGTAGAATGGGGGTGAC 84939
1073 1073

to bdqpefruitfly.berkeley.edu. All contigs in this submission meet the following cutoffs: length >= 200 bases.

* NOTE: This is a 'working draft' sequence. It currently consists of 142 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 536: contig of 456 bp in length
 * 457 536: gap of unknown length
 * 537 1004: contig of 468 bp in length
 * 1005 1084: gap of unknown length
 * 1085 1958: contig of 874 bp in length
 * 1959 2038: gap of unknown length
 * 2039 2248: contig of 210 bp in length
 * 2249 2328: gap of unknown length
 * 2329 2858: contig of 530 bp in length
 * 2859 2938: gap of unknown length
 * 2939 3325: contig of 387 bp in length
 * 3326 3405: gap of unknown length
 * 3406 4165: contig of 760 bp in length
 * 4166 4245: gap of unknown length
 * 4246 4716: contig of 471 bp in length
 * 4717 4796: gap of unknown length
 * 4797 5373: contig of 577 bp in length
 * 5374 5453: gap of unknown length
 * 5454 6050: contig of 597 bp in length
 * 6051 6130: gap of unknown length
 * 6131 6720: contig of 590 bp in length
 * 6721 6800: gap of unknown length
 * 6801 7461: contig of 661 bp in length
 * 7462 7541: gap of unknown length
 * 7542 8104: contig of 563 bp in length
 * 8105 8184: gap of unknown length
 * 8185 8505: contig of 321 bp in length
 * 8506 8585: gap of unknown length
 * 8586 9227: contig of 642 bp in length
 * 9228 9307: gap of unknown length
 * 9308 10091: contig of 784 bp in length
 * 10092 10171: gap of unknown length
 * 10172 11099: contig of 928 bp in length
 * 11100 11179: gap of unknown length
 * 11180 11747: contig of 568 bp in length
 * 11748 11827: gap of unknown length
 * 11828 12767: contig of 940 bp in length
 * 12768 12847: gap of unknown length
 * 12848 13398: contig of 551 bp in length
 * 13399 13478: gap of unknown length
 * 13479 14007: contig of 529 bp in length
 * 14008 14087: gap of unknown length
 * 14088 14867: contig of 780 bp in length
 * 14868 14947: gap of unknown length
 * 14948 15562: contig of 615 bp in length
 * 15563 15642: gap of unknown length
 * 15643 16859: contig of 1217 bp in length
 * 16860 16939: gap of unknown length
 * 16940 17986: contig of 1047 bp in length
 * 17987 18066: gap of unknown length
 * 18067 19341: contig of 1275 bp in length
 * 19342 19421: gap of unknown length
 * 19422 20926: contig of 1505 bp in length
 * 20927 21006: gap of unknown length
 * 21007 22153: contig of 1147 bp in length
 * 22154 22233: gap of unknown length
 * 22234 23236: contig of 1003 bp in length
 * 23237 23316: gap of unknown length
 * 23317 24607: contig of 1291 bp in length
 * 24608 24687: gap of unknown length
 * 24688 25620: contig of 933 bp in length
 * 25621 25700: gap of unknown length
 * 25701 26300: contig of 600 bp in length

26301 26380: gap of unknown length
 * 26381 27488: contig of 1108 bp in length
 * 27489 27568: gap of unknown length
 * 27569 28215: contig of 647 bp in length
 * 28216 28295: gap of unknown length
 * 28296 29440: contig of 1145 bp in length
 * 29441 29520: gap of unknown length
 * 29521 30587: contig of 1067 bp in length
 * 30588 30667: gap of unknown length
 * 30668 32042: contig of 1375 bp in length
 * 32043 32122: gap of unknown length
 * 32123 33740: contig of 1618 bp in length
 * 33741 33820: gap of unknown length
 * 33821 35375: contig of 1555 bp in length
 * 35376 35455: gap of unknown length
 * 35456 36258: contig of 803 bp in length
 * 36259 36338: gap of unknown length
 * 36339 38003: contig of 1665 bp in length
 * 38004 38083: gap of unknown length
 * 38084 39509: contig of 1426 bp in length
 * 39510 39589: gap of unknown length
 * 39590 40719: contig of 1130 bp in length
 * 40720 40799: gap of unknown length
 * 40800 42453: contig of 1654 bp in length
 * 42454 42533: gap of unknown length
 * 42534 43841: contig of 1308 bp in length
 * 43842 43921: gap of unknown length
 * 43922 45410: contig of 1489 bp in length
 * 45411 45490: gap of unknown length
 * 45491 46782: contig of 1292 bp in length
 * 46783 46862: gap of unknown length
 * 46863 48570: contig of 1708 bp in length
 * 48571 48650: gap of unknown length
 * 48651 51042: contig of 2392 bp in length
 * 51043 51122: gap of unknown length
 * 51123 52760: contig of 1638 bp in length
 * 52761 52840: gap of unknown length
 * 52841 54712: contig of 1872 bp in length
 * 54713 54792: gap of unknown length
 * 54793 56521: contig of 1729 bp in length
 * 56522 56601: gap of unknown length
 * 56602 58938: contig of 2337 bp in length
 * 58939 59018: gap of unknown length
 * 59019 60837: contig of 1819 bp in length
 * 60838 60917: gap of unknown length
 * 60918 62988: contig of 2071 bp in length
 * 62989 63068: gap of unknown length
 * 63069 65583: contig of 2515 bp in length
 * 65584 65663: gap of unknown length
 * 65664 68774: contig of 3111 bp in length
 * 68775 68854: gap of unknown length
 * 68855 71235: contig of 2381 bp in length
 * 71236 71315: gap of unknown length
 * 71316 74193: contig of 2798 bp in length
 * 74114 74194: gap of unknown length
 * 74194 77405: contig of 3212 bp in length
 * 77406 77485: gap of unknown length
 * 77486 80936: contig of 3451 bp in length
 * 80937 81016: gap of unknown length
 * 81017 83357: contig of 2341 bp in length
 * 83358 84337: gap of unknown length
 * 84338 87456: contig of 4019 bp in length
 * 87457 87532: gap of unknown length
 * 87537 91536: contig of 3996 bp in length
 * 91533 91612: gap of unknown length
 * 91613 97143: contig of 5531 bp in length
 * 97144 97223: gap of unknown length
 * 97224 105593: contig of 8370 bp in length
 * 105594 105673: gap of unknown length
 * 105674 112071: contig of 6398 bp in length
 * 112072 112151: gap of unknown length
 * 112152 121011: contig of 8860 bp in length
 * 121012 121091: gap of unknown length

```

* 121092 143113: contig of 22022 bp in length
* 143114 143193: gap of unknown length
* 143194 143587: contig of 394 bp in length
* 143588 143667: gap of unknown length
* 143668 143927: contig of 260 bp in length
* 143928 144007: gap of unknown length
* 144008 144842: contig of 835 bp in length
* 144843 144922: gap of unknown length
* 144923 145360: contig of 438 bp in length
* 145361 145440: gap of unknown length
* 145441 145915: contig of 475 bp in length
* 145916 145995: gap of unknown length
* 145996 146532: contig of 537 bp in length
* 146533 146612: gap of unknown length

alignment_scores:
  Quality: 1263.50      Length: 1512
  Ratio: 1.591          Gaps: 62
  Percent Similarity: 52.513      Percent Identity: 28.836

alignment_block:
US-09-294-298-4 x AC012162 ..
Align seg 1/1 to: AC012162 from: 1 to: 191504

11 ProValGlnValProCysProLeuLeuProThrAlaSer.....Le 24
112818 CCCCTACTAATGCCCATATCTTACGCTTCTTTCGACATAGCTCCAT 112867
112868 GAGAAAGCGTGGCGCGTGGATCAGCGCCACACCGCCCATTT..... 112910
24 uSerAlaAlaAlaLeu.....ProAlaAlaPheArgGlyA 37
112868 GAGAAAGCGTGGCGCGTGGATCAGCGCCACACCGCCCATTT..... 112910
37 snThrThrTpValAlaArgGlyArgValSerProGlyGlyLysGln 53
112911 .....TGGGAGCAACACAGCAGCAGCGGACACATGCCACCTCGC 112952
54 TyrSerMetGluAlaAlaProAlaAlaProphe..... 64
112953 GTTTCACCAACTTTTTCACAGAGTAGTTCCCGACATATATATATA 113002
65 .....ArgProSerGln.....GlyPheLeuSerA 73
113003 TATATATTCAGGTTCACATATTAATAATATGCCATTTCTTCTTGC 113052
73 rArgLeuLysSerSerIleLysArgThrLysSerGlnProLysLeuAsp 89
113053 ACAGA...TCCAATCCTTTGAAGCGGACCAAGTCGGTGACCAAGCTGGAG 113099
90 ArgThrSerSerPheArgGlnIleLeuProArgPheArgSerAlaAspHi 106
113100 CGGACCAAG..... 113108
106 sAspArgAlaArgLeuMetClnSerPheLysGluSerHisSerHisGluS 123
113109 .....CGCGGATCCGGCGGACTGAGGGGCTCCCGCTCGCACGAGA 113148
123 erLeuLeuSerProSerSerAlaAlaGluAla..... 133
113149 GTCTGCTGTCCAGTACGCGCTCATCTCCACCATAGGTGAGTTCTGGCTC 113198
133 ..... 133
113199 ATATGGAGTTTGGCCAAACCTCAACCATTTTCTTTCTTCTTCTTCT 113248
134 LeuGluLeuAsnLeuAspGluAspSerIleLysProValHisSers 150
113249 TCTAGATCTCTCGTACCTGGGCGGTGGGCGTGGCGCGCTGCATCAGT 113298
150 erIleLeuGlyGlnGluPheCysPheGluValThrThrSer....SerGly 165
113299 CGGTTCTGGGACGCTGCTACTGTTTCCAGGTACGGGGCGGCGCTCGTGGC 113348

```

```

166 ThrLysCysPheAlaCysArgSerAlaAlaGluArgAspLysTrpIleG1 182
113349 GAGCGGTACTACTCGATCGGATCGCGCAGGACGCGACCTTTGGATCTA 113398
182 uAsnLeuGlnArgAlaValLysProAsnLysAspAsnSerArgArgValA 199
113399 CTCGCTGGCGAAGTCGATCGCTCGAATGAGAGCACACGCTGCGCGG 113448
199 sAsnValLeuLysLeuTrpIleIleGluAlaArgGluLeuProProLys 215
113449 ACAACTCGCTGAAGATGTGGGTGTACGAGCGGAAATCTCGCGCCCAAG 113498
216 LysArgTyrTyrCysGluLeuLysLeuAspMetLeuTyrAlaArgTh 232
113499 AAGCGTTACTTTTGGAACTGCAATTTGACACAGACGCTGTACGGCGGAC 113548
232 rThrSerLysProArgSerAlaSerGlyAspThrValPheTrpGlyGluH 249
113549 TTCGGTGAAGCTGCAGACG.....GATCTGCTGTTTGGGGGGAGC 113589
249 isPheGluPheAsnAsnLeuProAlaValArgAlaLeuArgLeuHisLeu 265
113590 ACTTCGATTTCCCGACATACCGGAGATTAAATGTGATCACTGTTAACGTT 113639
266 TyrArgAspSerAspLysLysArgLysLysAspLysAlaGlyTyrValG1 282
113640 TTCGTGAGGTGGACAAAGAGAGCGGACCAATACCAATTTGTGGG 113689
282 yLeuValThrValProValAlaThrLeuAlaGlyArgHisPheThrGluG 299
113690 ATCGGTGAAGATACCGCTGCACGATGTCACCTGCAGATTCGCCGCGAAC 113739
299 InTrpTyrProValThrLeuProThrGlySerGlyGlySerGlyGlyMet 315
113740 AATGATATCCATA...CTGAGCGACAAAGCGGCGGACAGCTCGGCGAG 113786
316 GlySerGlyGlyGlyGlyGlySerGlyGlySerGlyGlySerGlyGlyLys 332
113787 ACATGGGCG.....GGCGGCGGAGTGGGTCCCAAGGACAA 113821
332 sGlyGlyCysProAlaValArgLeuLysAlaArgTyrGlnThrMetSerI 349
113822 GGAGCAATTCGCCGCTGAGGATCAAGTCTGTTTCCAGAGCAGCGACA 113871
349 leLeuProMetGluLeuTyrLysGluPheAlaGluTyrValThrAsnHis 365
113872 TCCTGCCCATCAATGTATACGGCAACTTTTGTGACGTACCTCAAGGAGA 113921
366 TyrArgMetLeuCysAlaValLeuGluProAlaLeuAsnValLysGlyLy 382
113922 TATAAGCGCGTGGCAGACCTGGAGCGGTGATCGGAGTCAAGGCCAA 113971
382 sGluGluValAlaSerAlaLeuValHisIleLeuGlnSerThrGlyLysA 399
113972 GGAGACATTCGACAGGCACTGGTGTGCTGATGTCACGACGAGGATGG 114021
399 laLysAspPheLeuSerAspMetAlaMetSerGluValAspArgPheMet 415
114022 CGGGCGCGCTTCCTCACCGATGTGGTGGCCCTCGATCTGCTGCGAGTTGC 114071
416 GluArgGluHisLeuIlePheArgGluAsnThrLeuAlaThrLysAlaI1 432
114072 GAT...CAGAGGCTTACGAGGGCACTCTTGGCCACCAAGAGCAT 114118
432 eGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuLysAspAlaIleG 449
114119 GGAGCATTCCTCAAGCTGACGGCGCAACAGTATCTGACGAGCACACTAT 114168
449 yGluPheIleArgAlaLeuTyrGluSerGluGluAsnCysGluValAsp 465
114169 CGGCACCAATAAAGAGCTAATTCAGTCGAGAGGAGTCCGAGGTGGAT 114218
466 ProLeLys.....CysThrAlaSerSerLeuAlaGluHisGlnAl 479

```

```

114219 CCCACCAAGACGAGCGGTTCGCGCGGGTTCCTGCAGCAGCAGCGC 114268
479 aAsnLeuArgMetCysCysGluLeuAlaLeuCysLysValValAsnSerH 496
114269 CGCCTTGGCTGGCGCGGTTCGCGAGGGCGGTGGCAGTGCATCTCGAATCGC 114318
496 iScysValPheProArgGluLeuLysGluValPheAlaSerTrpArgLeu 512
114319 ACAAGCATTTCCCGCCAGTTCAGAAATTCCTTCGCGAGTTCGCGGAG 114368
513 ArgCysAlaGluArgGluArgGluAspIleAlaAspArgLeuSerAl 529
114369 CGCTTGCAGCAGCTGGCGCGGTTCAGGATATGGCCGCAACCTGATCTCGC 114418
529 aSerLeuPheLeuArgPheLeuCysProAlaIleMetSerProSerLeuP 546
114419 GAGCATTTTCCTGGCTTCCTGTGCGCGGCATCTGTGCGCGTCTCT 114468
546 heGlyLeuMetGlnGluTrpProAspGluGlnThrSerArgThrLeuThr 562
114469 TCAATATCACCAGGGAATCTGCGTCCGACGCGGCCACCGCAATCTCACA 114518
563 LeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerLysPheThrSe 579
114519 CTGGTGGCAAGACCTGCAAACTTTCGCAATTTCCACCGCTTCCAGGG 114568
579 rLysGluAspPheLeuGlyPheMetAsnGluPheLeuGluLeuTrpG 596
114569 CAAGAGAACTTTATGGAGTTCTCAACGATTTCTCGAGCAGGAGCGG 114618
596 lysSerMetGlnGlnPheLeuTyrrGluIleSerAsnLeu 608
114619 CTCGCATCAACAGTTTCTGGAGATTATATCCACGCGCGCGAGCACCCA 114668
609 .....AspThrLeuThrAsnSerSerPheGluGlyTyrrIleAspLe 623
114669 GCTCCAGACTCGATCTCGAT.....TGGCGCGGTTCATCGACCA 114709
623 uGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuProG 640
114710 GGGCAACACATGTCTTCCATACATACAGTTTGTCTCAGGAAAGCTTGCCA 114759
640 InLeuSerLysGluAlaLeuLysLeuGlyProLeuProArgLeuLeu 656
114760 AGCTGCCGAGGCCAGCAGCAGCGTGGATCCGTTGCAGCATATCTC 114809
657 SerAspIleSer.....ThrAlaLeu..... 663
114810 GATGAATACGCCGAGCAAAAGACATGGCATGGGCACAGCATGCGCGG 114859
663 ..... 663
114860 TGGATATTTTCGGGCCACCTCGTCTAGCACATGATAGCAGCGGAATC 114909
664 .....ArgAsnPro.....AsnIle 668
114910 AGGAGAAATCGCAATCGGGATCTCGGCTCGCAGCGTGCCTCCACTCG 114959
669 GlnArgGlnProSerArgGlnSerGluArgAlaArgSerGlnProMetVa 685
114960 GAGCAGTTACTGCCACAACAAGAGCGAGTTGGCCCGCGCGCATGCGGAT 115009
685 lLeuArgGlyPro.....SerAlaGlu.....MetGlnGly.TyrMet 697
115010 TGTTAGTAACCAATTTATCTCGGAGCGCGGCATCATGCGAGGAGTACTTA 115059
698 MetArgAspLeuAsnSerSerIleAspLeuGlnSerPheMetAlaArgG 714
115060 CGCCGAAATTCCTGGAGAGAGATATCTTTAGATACATGATCCACCGGT 115109
714 yLeuAsnSerSerMetAspMetAlaArgLeuProSerProThrLysGluL 731

```

```

115110 AATG..... 115113
731 ysProProProProGlyGlyGlyAspLeuPheTyrValSer 747
115113 ..... 115113
748 ArgProProLeuAlaArgSerSerProAlaTyrrCysThrSerSerSerAs 764
115114 .....GCTTACTGCAGCAGCAGCAGCAG 115136
764 pIleThrGluProGluGlnLysMetLeuSerValAsnLysSerValSerM 781
115137 CAGCAGAGC.....AGCAGCAACAACAGCAACAGCA 115168
781 tLeuAspLeuGlnGlnLysAspGlyProGlyArgLeuAsnSerSerSer 797
115169 ACATG.....AGCAGCTGCAACGACATGGCC 115194
798 ValSerAsn.....LeuAlaAlaValGly..AspLeuLeuHisSerSer 811
115195 ATCAGCAACAGCCGCCACACAGCATCCACTCCAGATGCTCTCCAATCA 115244
812 GlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaProAlaGlyAr 828
115245 CAACCTCCATT.....GCCGCAACCAATATATAGTTCCGCCAGGAGGCT 115291
828 gLeuSerGlnGlySerGlySerSerIleThrAlaAlaGlyMetArgLeu 845
115292 GCAGCATGCCCAATCGCAGACCTCGATGGCTCTCATCGCTT..... 115334
845 erGlnMetGlyValThrAspGlyValProAlaGlnGlnLeuArgIle 861
115334 ..... 115334
862 ProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaAspGlyProGl 878
115334 ..... 115334
878 yProProAlaGlyHisGlyGlySerSerGly.....HisGlyProProS 893
115335 .....AATGGGAGCAGCAGCAATTTGCTGCAGCGC..... 115364
893 erSerHisHisHisHisHisHis.....HisHisHisArg 905
115365 .....CACCAGCAGCATGCCCATCACCCGCACTGCATCCACATCAC 115409
906 GlyGlyGluProProGlyAspThrPheAlaProPheHisGlyTyrSerly 922
115410 TGC.....CGCCG..... 115418
922 sSerGluAspLeuSerThrGlyValProLysProProAlaAlaSerIleL 939
115419 .....GCCACACAGCAGTGCCTCCAGCACA 115446
939 euHis.....SerHisSerTyrSerAspGluPheGlyPro 950
115447 TGGAGCGCATGGATCGATGAATATCCGTATATGTCCGAC..... 115487
951 SerGlyThrAspPheThrArgArgGlnLeuSerLeuGlnAspAsnLeuGl 967
115488 AATGCAATGACTAGCAG..... 115505
967 nHisMetLeuSerProGlnIleThrIleGlyProGlnArgProAlaP 984
115506 .....ACCAGCAGCGCTTCGAGCAGCTCCAGCAGACTGCCACGGA 115548
984 roSerGlyPro.....GlyGlyGlySerGlyGlySerGlyGlyGly 998
115549 ATGGAATCCCAATGCCAATGCCAAGCTGGGAGCAGCAACAATAACAG 115598
999 GlyGlyGlyGlnProProProLeuGlnArgGlyLysSerGlnGlnLeu. 1014
115599 AGCGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 115648

```

1015ThrValSerAlaAlaGlnLysProArgPro..... 1024
 115649 GTTCGATACGAGCAGTGTCTGCTCAAGTGCAGCCACACCCATGATGA 115698
 1025Ser.SerGlyAsnLeuLeuGlnSerProGluProSerT 1037
 115699 ATTCAGTTTGGAGCAGCGGGGGCCGAGAGTATCAACAACATGAAT 115748
 1037 yGlyProAlaArgProArg.....Gln 1044
 115749 TTGGGAATACCCGATCACTACAGTGGGTATGCGGGGTAATAATTGAA 115798
 1045 GlnSerLeuSerLysGluGlySerIleGlySerGlyGlySer..... 1059
 115799 TCCCAACTCGAATATGCCAAGAACTTGGAGGATCTGGAGATCTGTTC 115848
 1060GlyGlyGlyGlyGlyGlyGlyGlyLeuLysProSerIleThrL 1073
 115849 AGTACGCCGAGGAGCATGACGTGGCGAACCAGCAGCAACCATCAATCAC 115898
 1073 ysGlnHisSer...GlnThrProSerThr.IeuAsnPro...ThrMetPr 1087
 115899 AACCAGGTCACAGAACCCAGGAGACATCTGAAGCCGCGCCGCTTCC 115948
 1087 oAlaSerGluArgThrValAlaTrpValSerAsnMetProHisLeuSerA 1104
 115949 CGCAAGGAGCAGTGTGGCGGAAAGCAGT.....CACTGCAGTT 115989
 1104 laAspileGluSerAlaHisIleGluArgGluGluThrLysLeuLysGlu 1120
 115990 CTGGTACCAGGATCTCCACAATCCG..... 116018
 1121 TyrSerLysSerMetAspGluSerArgLeuAspArgValLysGluTyrGl 1137
 116019 ...TGCCCTCCAGTCTCCAGTCCCGTCCGAG..... 116048
 1137 uGluGluIleHisSerLeuLysGluArgLeuHisMetSerAsnArgLysL 1154
 116049AGCCAGCTGAAGGCGCGGATGGGAGTCAATATCGCCGCG 116088
 1154 euGluGluTyrGluArgLeuLeuSerGlnGluGluGlnThrSerLys 1170
 116089 TG...GCCTTCAAGAAATCCCTCTATCATGCTTCAGCCCAACATGCTCG 116135
 1171 IleLeuMetGlnTyrGlnAlaArgLeuGluGlnSerGluLysArgLeuAr 1187
 116136 TCCAGATCATCGCCAGAGTAAATACACACAGCAGCAGCAGCAGCA 116185
 1187 gGlnGlnGlnValGluLysAspSerGlnIleLysSerIleIleGlyArgL 1204
 116186 ACAACAACAGTTT.....GGCAGCGCTTGAACAACCAATTTGGCGCGGAC 116229
 1204 euMetLeuValGluGluLeuArgArgAspHisProAlaMet.AlaGl 1220
 116230 TGGTGGCGCGAGGCGCTTCTCAACAGTGGCGGAGCTTGGAGGGG 116279
 1220 uProLeu.....ProGluProLys..... 1226
 116280 GCCACTTTGACGCCCGAGCTCTCGGAGCAACAGCTCTCGGCGGATAATTA 116329
 1227LysArgLeuLeuAspAlaGlnArgGlySerPheProPro 1239
 116330 CTTTCAAGTATGACGGCTGAGCTGTGGAGCGGGTATTGCCACCAAT 116379
 1240 TrpValGlnGlnThrArgValArgTrpProHisLeuGlyThrAlaTrpPr 1256
 116380 TGGAGGCTCAACGCTCGC.....TCAGCGGCGGAGTAGTCTCTCT 116420
 1256 oProGln.....ProHisP 1261
 116421 ACCTCAGCATCTGGCTCCACCTCGAATCTTGGCAAGAGCGGCGCTCATC 116470

1261 roProHisProGlyCysArgSerGlnArgThrAlaSerSerGlyThrPro 1277
 116471 CGCCTACGGCGCGCTGAATGGCCGCTGAAGCGCAGGATGCTCTACGGCA 116520
 1278 GlnThrThrSerProPro.....SerIleThrAspLeuLeuPr 1290
 116521 GTGGCTACCGCGCGCGAGTGAATAATGGGCTATGGCTGTTCACACTTC 116570
 1290 o.Cys.....AlaProTyrProProProSerValThrAspLeuLeu 1304
 116571 AGTGGCGCGGAGACATCAACATCCCAACAGCAGCAGCAATCGCAT 116620
 1305 ProSerAlaProAspProGlyThrSerProThrThrArgThr..... 1318
 116621 GCAGCAGCAGCAGCAAGGAGGATCGGATCAGCAACACAGCAGTATGCCG 116670
 1319GlyArgHisGlnGlyThrAlaGlyLeuSerProLeuThrProp 1333
 116671 GCAGTGTGGCGCGCAGCTGGCCATCATCAGCGGCTCAGAGGCGC 116720
 1333 ro.....TrpGlyThrHis 1337
 116721 CTGAGCTTGGACTCGCGCGCAC 116743
 seq_name: gb_inl:AE003506
 seq_documentation_block:
 LOCUS AE003506 300994 bp DNA INV 04-OCT-2000
 DEFINITION Drosophila melanogaster genomic scaffold 14200001386053 section 23
 of 30, complete sequence.
 ACCESSION AE003506 AE002593
 VERSION AE003506.1 GI:7293355
 KEYWORDS HTS
 SOURCE fruit fly.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 Muscomorpha; Ephydroidea; Prosoiphilidae; Drosophila.
 REFERENCE
 1 (bases 1 to 300994)
 AUTHORS
 Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
 Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F.,
 George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
 Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X.,
 Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
 Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
 Mikos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
 Andrews-Brannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
 Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
 Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
 Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
 Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
 Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
 Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
 Dodson,K., Doup,I.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
 Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferrieria,S.,
 Feilichmann,W., Fosler,C., Gabriellian,E., Garg,N.S.,
 Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z.,
 Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
 Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
 Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z.,
 Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
 Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
 Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C.,
 McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
 Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
 Muzny,D.M., Nelson,D.B., Nelson,D.R., Nelson,K.A., Nixon,K.,
 Nusskern,D.R., Pacle,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
 Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
 Saunders,R.D., Scheefer,F., Shen,H., Shue,B.C., Siden-Kiamos,I.,
 Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
 Stapleton,M., Strong,R., Sun,E., Swirskas,R., Tector,C., Turner,R.,
 Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassarman,D.A.,
 Weinstock,G.M., Weissbach,J., Williams,S.M., Woodage,T.,

* 590 1201: contig of 612 bp in length
 * 1202 1281: gap of unknown length
 * 1282 2394: contig of 1113 bp in length
 * 2395 2474: gap of unknown length
 * 2475 3512: contig of 1038 bp in length
 * 3513 3593: gap of unknown length
 * 3593 4500: contig of 908 bp in length
 * 4501 4580: gap of unknown length
 * 4581 5466: contig of 885 bp in length
 * 5466 5545: gap of unknown length
 * 5546 6253: contig of 708 bp in length
 * 6254 6333: gap of unknown length
 * 6334 7105: contig of 772 bp in length
 * 7106 7185: gap of unknown length
 * 7186 8077: contig of 892 bp in length
 * 8078 8157: gap of unknown length
 * 8158 9047: contig of 890 bp in length
 * 9048 9127: gap of unknown length
 * 9128 9530: contig of 403 bp in length
 * 9531 9610: gap of unknown length
 * 9611 10173: contig of 563 bp in length
 * 10174 10253: gap of unknown length
 * 10254 10841: contig of 588 bp in length
 * 10842 10921: gap of unknown length
 * 10922 11807: contig of 886 bp in length
 * 11808 11887: gap of unknown length
 * 11888 13165: contig of 1278 bp in length
 * 13166 13245: gap of unknown length
 * 13246 14526: contig of 1281 bp in length
 * 14527 14606: gap of unknown length
 * 14607 15801: contig of 1195 bp in length
 * 15802 15881: gap of unknown length
 * 15882 16640: contig of 759 bp in length
 * 16641 16720: gap of unknown length
 * 16721 17679: contig of 959 bp in length
 * 17680 17759: gap of unknown length
 * 17760 18797: contig of 1038 bp in length
 * 18798 18877: gap of unknown length
 * 18878 19955: contig of 1078 bp in length
 * 19956 20335: gap of unknown length
 * 20336 20931: contig of 896 bp in length
 * 20932 21011: gap of unknown length
 * 21012 21643: contig of 632 bp in length
 * 21644 21723: gap of unknown length
 * 21724 23416: contig of 1693 bp in length
 * 23417 23496: gap of unknown length
 * 23497 24457: contig of 961 bp in length
 * 24458 24537: gap of unknown length
 * 24538 25955: contig of 1418 bp in length
 * 25956 26035: gap of unknown length
 * 26036 27838: contig of 1803 bp in length
 * 27839 27918: gap of unknown length
 * 27919 29686: contig of 1768 bp in length
 * 29687 29766: gap of unknown length
 * 29767 30976: contig of 1210 bp in length
 * 30977 31056: gap of unknown length
 * 31057 32460: contig of 1404 bp in length
 * 32461 32540: gap of unknown length
 * 32541 34644: contig of 2104 bp in length
 * 34645 34724: gap of unknown length
 * 34725 36341: contig of 1617 bp in length
 * 36342 36421: gap of unknown length
 * 36422 37864: contig of 1443 bp in length
 * 37865 37944: gap of unknown length
 * 37945 39424: contig of 1480 bp in length
 * 39425 39504: gap of unknown length
 * 39505 41341: contig of 1837 bp in length
 * 41342 41421: gap of unknown length
 * 41422 42970: contig of 1549 bp in length
 * 42971 43051: gap of unknown length
 * 43051 44419: contig of 1369 bp in length
 * 44420 44499: gap of unknown length
 * 44500 46484: contig of 1985 bp in length

* 46485 46564: gap of unknown length
 * 46565 48446: contig of 1882 bp in length
 * 48447 48526: gap of unknown length
 * 48527 50693: contig of 2167 bp in length
 * 50694 50773: gap of unknown length
 * 50774 53578: contig of 2805 bp in length
 * 53579 53658: gap of unknown length
 * 53659 55753: contig of 2095 bp in length
 * 55754 55833: gap of unknown length
 * 55834 58457: contig of 2624 bp in length
 * 58458 58537: gap of unknown length
 * 58538 60171: contig of 1634 bp in length
 * 60172 60251: gap of unknown length
 * 60252 63621: contig of 3370 bp in length
 * 63622 63701: gap of unknown length
 * 63702 66219: contig of 2518 bp in length
 * 66220 66299: gap of unknown length
 * 66300 69074: contig of 2775 bp in length
 * 69075 69154: gap of unknown length
 * 69155 72905: contig of 3751 bp in length
 * 72906 72985: gap of unknown length
 * 72986 75963: contig of 2978 bp in length
 * 75964 76043: gap of unknown length
 * 76044 78866: contig of 2823 bp in length
 * 78867 78946: gap of unknown length
 * 78947 82045: contig of 3099 bp in length
 * 82046 82125: gap of unknown length
 * 82126 85135: contig of 3010 bp in length
 * 85136 85215: gap of unknown length
 * 85216 88550: contig of 3335 bp in length
 * 88551 88630: gap of unknown length
 * 88631 92611: contig of 3981 bp in length
 * 92612 92691: gap of unknown length
 * 92692 98789: contig of 6018 bp in length
 * 98790 98789: gap of unknown length
 * 98791 104147: contig of 5358 bp in length
 * 104148 104227: gap of unknown length
 * 104228 111796: contig of 7569 bp in length
 * 111797 111876: gap of unknown length
 * 111877 118488: contig of 6612 bp in length
 * 118489 118569: gap of unknown length
 * 118569 126924: contig of 8356 bp in length
 * 126925 127004: gap of unknown length
 * 127005 137051: contig of 10047 bp in length
 * 137052 137131: gap of unknown length
 * 137132 148859: contig of 11728 bp in length
 * 148860 148939: gap of unknown length
 * 148940 149554: contig of 615 bp in length
 * 149555 149634: gap of unknown length
 * 149635 150197: contig of 563 bp in length
 * 150198 150277: gap of unknown length
 * 150278 150907: contig of 630 bp in length
 * 150908 150987: gap of unknown length
 * 150988 151646: contig of 659 bp in length
 * 151647 151726: gap of unknown length
 * 151727 152074: contig of 348 bp in length
 * 152075 152154: gap of unknown length
 * 152155 152902: contig of 748 bp in length
 * 152903 152982: gap of unknown length
 * 152983 153624: contig of 642 bp in length
 * 153625 153705: gap of unknown length
 * 153706 154242: contig of 538 bp in length
 * 154243 154322: gap of unknown length
 * 154323 154739: contig of 417 bp in length
 * 154740 154819: gap of unknown length
 * 154820 155338: contig of 519 bp in length
 * 155339 155418: gap of unknown length
 * 155419 155993: contig of 575 bp in length
 * 155994 156073: gap of unknown length
 * 156074 156806: contig of 733 bp in length
 * 156807 156886: gap of unknown length
 * 156887 157541: contig of 655 bp in length
 * 157542 157621: gap of unknown length

```

* 157622 158208: contig of 587 bp in length
* 158209 158288: gap of unknown length

alignment_scores:
  Quality: 1257.50      Length: 1432
  Ratio: 1.642          Gaps: 57
  Percent Similarity: 53.492  Percent Identity: 29.190

alignment_block:
US-09-294-298-4 x AC012161/rev ..
Align seg 1/1 to reverse of: AC012161 from: 1 to: 202741

70 PheLeuSerArgArgLeuLysSerSerIleLysArgThrLysSerGlnPr 86
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126785 TTCTTTGCCACAGA...TCCAACTCTTTGAAGCGGACCAAGTGGTGAC 126739

86 oLysLeuAspArgThrSerSerPheArgGlnIleLeuProArgPheArgS 103
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126738 CAAGCTGGAGCGGACCAAG..... 126720

103 erAlaaspHisaspArgAlaArgLeuMetGlnSerPheLysGluSerHis 119
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126719 .....CGCGGATCCGGCGGACTGAGGGGCTCCCGC 126690

120 SerHisGluSerLeuLeuSerProSerSerAlaAlaGluAla..... 133
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126689 TCGCAGGAGAGTCTGCTGTCAGAGTACACCGCGTCATGTCACCAATAGTGA 126640

133 ..... 133

126639 GTTCTGCTCATATGAGGAGTTTGGCCAAACCCCTCAAAGCAATTTCTTTT 126590

134 .....LeuGluLeuAsnLeuAspGluAspSerIleIleLysPro 146
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
12589 TCTTTCTTCTTAGATCTCTCGTCACATGGGCGGTGGCGGTGGCGGCC 126540

147 ValHisSerSerIleLeuGlnGluPheCysPheGluValThrThrse 163
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126539 GTGCATCAGTCGGTCTGGGAGCTGCTGCTCCTGTTCCAGGTACGGGCGG 126490

163 r...SerGlyThrLysCysPheAlaCysArgSerAlaAlaGluArgAspL 179
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126489 GCCTGTGGCGGCGGTACTACTATCGGATCGCGCGGAGGCGGCGAC 126440

179 ysrpIleGluAsnLeuGlnArgAlaValLysProAsnLysAspAsnSer 195
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126439 TTTGGATCTACTCGCTGCGCAAGTCGATCGCTCCGAATGCGAGCACACG 126390

196 ArgArgValAspAsnValLeuLysLeuTrpIleIleGluAlaArgGluLe 212
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126389 CGTGCAGCGGACCACTCGCTGAAGATGTGGGTGTACGAGGCGGAAATCT 126340

212 uProProLysLysArgTyrTyrCysGluLeuLysCysLeuAspAspMetLeu 229
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126339 GCGGCCCAAGACGTTACTTTTGGCACTGCAATTTGACAAAGACGCTGT 126290

229 yrlaArgThrThrSerLysProArgSerAlaSerGlyAspThrValPhe 245
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126289 ACGCGCGGACTTCGGTGAAGTCGACAGC.....GATCTGCTGTTT 126249

246 TrpGlyGluHisPheGluPheAsnAsnLeuProAlaValArgAlaLeuAr 262
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126248 TGGGGGGAGCACTTCGATTTCCCGGACATACCGGATTAATGTATCATC 126199

262 leuHisLeuLysLeuArgSerAspLysArgLysAspLysAlaG 279
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126198 TGTATACGTTTCCGTGAGTGGCAGACAGACGAGCGGACCAATACC 126149

279 lyPyrValGlyLeuValThrValProValAlaThrLeuAlaGlyArgHis 295
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126148 AATTTTGGGATCGGTGAAGATACCGGTGCACGATGTCACTCCAGATTG 126099

```

```

296 PheThrGluGlnTrpTyrProValThrLeuProThrGlySerGlyGlyse 312
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126098 CCCTCGCAACAATGGTATCCATA...CTGACGACAGAGCGGGGACAG 126052

312 rGlyGlyMetGlySerGlyGlyGlyGlySerGlyGlySerGlyGly 329
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126051 TCTGGGAGGACATCGGC.....GGCGGCGGACGTGGGT 126017

329 lyLysGlyLysGlyGlyCysProAlaValArgLeuLysAlaArgTyrGln 345
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126016 CCAAGGACAAAGAGCAATTCGCCACGCTGAGGATCAAGTGTCTTCCAG 125967

346 ThrMetSerIleLeuProMetGluLeuTyrLysGluPheAlaGluTyrVa 362
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125966 AGCACCCGACATCTGCCATCANTGTAGCGCACTTTTTCACGTACT 125917

362 lThrAsnHisTyrArgMetLeuLysAlaValLeuGluProAlaLeuAsnV 379
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125916 CAAGGAGAACTATAAGCGCGTGTGCGAGACCTTGGAGCGGTGATCGGAG 125867

379 allysGlyLysGluValAlaSerAlaLeuValHisIleLeuGlnSer 395
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125866 TCAAGGCCAAGGAGGACATTTGGACAGGCTGGTGTCTGATGCACGCA 125817

396 ThrGlyLysAlaLysaspPheLeuSerAspMetAlaMetSerGluValas 412
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125816 CAGGATTTGGCGGCGCTTCTCACCAGATGTGGTGGCCCTCGATCTGCT 125767

412 pArgPheMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAla 429
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125766 GCGAGTTGGCGAT...CAGAGGCTTACGTTACAGGGGCAACTCTTGGCCA 125720

429 hrLysAlaIleGluGluTyrMetArgLeuIleGlnLysTyrLeuLys 445
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125719 CCAGAGCATGAGGACATCTCTCAGCTGACGGCGGACGATCTGCGAG 125670

446 AspAlaIleGlyGluPheIleArgAlaLeuTyrGluSerGluGluAsnCy 462
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125659 GACACATATCCGCACCAATAAAGAGCTAATTCAGTCGAGAGGGGACTG 125620

462 sGluValAspProIleLys.....CysThrAlaSerSerLeuAlaG 476
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125619 CGAGTGGATCCCAACCAAGAGCGGTTCGTCGGCGGTTCTGTCGACC 125570

476 luHisGlnAlaAsnLeuArgMetCysGluLeuAlaLeuLysVal 492
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125569 GACAGAGCGCGCTTGGTGGCGCGTCCGAGGGCGTGGCAGTGCATC 125520

493 ValAsnSerHisCysValPheProArgGluLeuLysGluValPheAlase 509
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125519 TTCGATCGCACCAAGCATTTCCCGCCAGTTACGAAATTTGCTTTGCGAC 125470

509 rTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaAspArgL 526
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125459 GTTCCGGAGCGCTTGCAGCAGCGCTGCGGATAGGATATGGCGCAACC 125420

526 euIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetSer 542
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125419 TGATCTCGGCGAGCATTTTCTCGGCTTCTGTCGCCGCCCATCTGTCG 125370

543 ProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnThrSerAr 559
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125369 CCGTCGCTCTCAATATACCAAGCAACTGCCGTCGCGACAGCGGCCACCG 125320

559 gThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerL 576
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125319 CAATCTCACACTGGTGGCCCAAGACCTGCNAACATTTGGCCATTTCCACC 125270

576 yspThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGlu 592
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125269 GCTTCCAGGCAAGAGAGACTTTTATGAGTTTCTCAACGATTTCTCTCGAG 125220

593 LeuGluTrpGlySerMetGlnGlnPheLeuGluIleSerAsnLeu... 608

```

```
125219 CAGGAGCCGCTCGCATGCAACAGTTTCTGGAGATTATATCCAGCGGCC 125170
125169 GGAGCACCAGCTCCAGACTCGATCCTCGAT .....TGGCCCGGT 125129
609 .....AspThrLeuThrAsnSerSerPheGluGlyT 620
125169 GGAGCACCAGCTCCAGACTCGATCCTCGAT .....TGGCCCGGT 125129
620 yTleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGlu 636
125128 ACATCGACCAAGCAACAGTTTCTCATACATACACAGTTTCTCAGCGAA 125079
637 ValLeuProGlnLeuSerLysGluAlaLeuLeuLysLeuGlyProLeuPr 653
125078 AGCTGCGCAAGCTCGCGGAGCCAGGACGACGAGCTGGATCGTTGCA 125029
653 oArgLeuLeuSerAspIleSer.....ThrA 662
125028 GCATATTCTCGATGAATCAGCGGACCAAGAGCATGGCATGGGCACAG 124979
662 laLeu..... 663
124978 CACTGCGCGGTGGATATTTCGCGGCCACCTCGTCTACGCACTCGATAGCC 124929
664 .....ArgAsnPro..... 666
124928 AGCGAGAAATCAGGAGAAATCGCAATCCGGATCCTCGGGCTCGCAGCTGG 124879
667 ....AsnIleGlnArgGlnProSerArgGlnSerGluArgAlaArgSerG 682
124878 CTCCAACCTCGGAGCACTTACTGCCCAACAAGCCAGTTGGCCAGCGCG 124829
682 lnProMetValLeuArgGlyPro....SerAlaGlu.....MetGln 694
124828 AGCATGGGATTTAGTAAACCATTTATCTCGGAGCGCGGCATCATCGCA 124779
695 Gly.TyrMetMetArgAspLeuAsnSerSerIleAspLeuGlnSerPheM 711
124778 GGAGTACTTACGCGGAATCTCTCGGAGAGAAATATCTTTAGATACAATGA 124729
711 etAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSerPro 727
124728 TCCACGCGTTAATG..... 124715
728 ThrLysGluLysProProProProProGlyGlyGlyLysAspLeuPh 744
124715 ..... 124715
744 eTyrValSerArgProProLeuAlaArgSerSerProAlaTyrCysThrS 761
124714 .....GCTTACTCGACA 124702
761 erSerSerAspIleThrGluProGluGlnLysMetLeuSerValAsnLys 777
124701 GCAGCAGCAGCAGCAGCAGC.....AGCAGCAACAAC 124670
778 SerValSerMetLeuAspLeuGlnGlyAspGlyProGlyGlyArgLeuAs 794
124669 AGCAACAGCAACATC.....AGCAGCTGCAC 124644
794 nSerSerSerValSerAsn.....LeuAlaAlaValGly..AspLeuLe 808
124643 CAGCATGGCCATCAGCAACAGCGGACCAAGCATCCAGATGCT 124594
808 uHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaP 825
124593 CTCCAATTCACAAACCTCCCAAT...GCCGGCAACCAATATATGAGTTGCG 124547
825 roAlaGlyArgLeuSerGlnGlySerSerIleThrAlaAlaGly 841
124546 CAGGAGGCTCCAGCATGCCCAATCGACACCTCGATGCGTCTCATCG 124497
842 MetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGlnG 858
124496 CTT..... 124494
858 nLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAla 875
124494 ..... 124494
875 spGlyProGlyProProAlaGlyHisGlySerSerGly.....His 889
124493 .....AATGGGAGAGCAGCAATTTGCTGTGCAC 124467
890 GlyProProSerSerHisHisHisHisHisHis.....HI 902
124466 GGC.....CACCAGCAGCATGCCATCACCAGCAACATGCA 124429
902 sHisHisArgGlyGluProProGlyAspThrPheAlaProPheHisG 919
124428 TCCACATCAGTC.....CCGCGG..... 124410
919 lyTyrSerLysSerGluAspLeuSerThrGlyValProLysProProAla 935
124409 .....GCGCCACAGCAGTGC 124392
936 AlaSerIleLeuHis.....SerHisSerTyrSerAspG 947
124391 TCCAGCACTATGGAGCCGATGGATCGATGCACTATCCGTATATGCGCA 124342
947 uPheGlyProSerGlyThrAspPheThrArgGlnLeuSerLeuGlnA 964
124341 C.....AATGGCAATGACTACGAG..... 124323
964 spAsnLeuGlnHisMetLeuSerProProGlnIleThrIleGlyProGln 980
124322 .....ACCAGCAGCCTTCGAGCACTCTCTCCAGGACA 124290
981 ArgProAlaProSerGlyPro.....GlyGlyGlySerGlyGlyGlySe 995
124289 CTGCCACGGAATGGAAATCCCAATGCCAATGCCAAGTGGCAGCAGCAA 124240
995 rGlyGlyGlyGlyGlyGlnProProPro.LeuGlnArgGlyLysSer 1011
124239 CAATACACAGAGCGGAGTACGATGACATGCACGGGGAGTTCCAANTCC 124190
1012 GlnGlnLeu.....ThrValSerAlaAlaGlnLysProAr 1023
124189 AGATCTCTGGTTCGATACGAGCAGTCTTTTGTCTGCAAGTCGCCACA 124140
1023 gPro.....Ser.GlyAsnLeuLeuGlnSerPro 1033
124139 CCATGATGAATCCAGTTTGGGACAGCGGGCGCGGAGAGTCTATCA 124090
1034 GluProSerTyrGlyProAlaArgProArg..... 1043
124089 CAAACTGAATTTGGGAATACCCGATCATCAGGTGGCTATGTGCGGGGTA 124040
1044 .....GlnGlnSerLeuSerLysGluGlySerIleGlyGlySerGlyG 1058
124039 ATAAATTTGAATCCCACTCGAATATGCCAAGAATCTGGAGGATCTGGAC 123990
1058 lySer.....GlyGlyGlyGlyGlyGlyLeuLysPro 1069
123989 GATCTGTTCAGTACGCGGAGGAGTACGATGGCGGAACAGCAACCA 123940
1070 SerIleThrLysGlnHisSer...GlnThrProSerThr.LeuAsnPro. 1084
123939 TCACAAATCACAACAGGTCAGCAGAACACAGGACATCTCAAGCCCG 123890
1085 ..ThrMetProAlaSerGluArgThrValAlaTrpValSerAsnMetPro 1100
123889 CGCGCTTCCGCGCAAGGAGCAGCTGTGCGGCAAAAGCAGT..... 123849
1101 HisLeuSerAlaAspIleGluSerAlaHisIleGluArgGluGlyLys 1117
123848 CACTCGAGTTCTGGTACCAGAGCATCTCCACAARTCCC..... 123810
```


408 etSerGluValAspArgPheMetGluArgGluHisLeuIlePheArgGlu 424
1985 CTGCGGAGCGTGGCGGCTGTGGAGCGCGCTGAGGCGCTGCTGTCCGGAA 2034
425 AsnThrLeuAlaThrLysAlaIleGluGlyTrpMetArgLeuIleGlyG1 441
2035 AACACATTGGCCACCAAGGCTATCGATGAGTACATGAAGCTCGTGGCACA 2084
441 nlyStrLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTyrcLus 458
2085 GGATTACTCCAGGAGACCCGTGGACAGGTGTGCGGCGTCTCTGTGCTT 2134
458 erGluGluAsnCysGluValAspProIleLysCysThrAlaSerSerLeu 474
2135 CTACTGAGGAGCTGTGAAGTGGACGCCACGAAATGTCAGAGCTCGGAGCTG 2184
475 AlaGluHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuTyrcLys 491
2185 CCAGAGCACCAGGCCAGACTTCGGAACAGCTGCGAGGAGGTCTCGAAAC 2234
491 sValValAsnSerHisCysValPheProArgGluLeuLysGluValPheA 508
2235 CATTAATCAATCTACGACTGGTTCCTCCGAGCTGGGATCGGTGTCT 2284
508 laSerTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaAsp 524
2285 CAAGCTGGCGAAGACGTAAAGAACCTGGCTGAGGTGCTGGGCGCC 2334
525 ArgLeuIleSerAlaSerLeuPheLeuArgPheLeuLeuTyrcLysProAlaIleMe 541
2335 CGACTGGGTGCGCTCCCTCTCTCTCGGCTCTGTGCTGCTGCCATCCT 2384
541 tSerProSerLeuPheGlyLeuMetGlnGluTyTrpAspGluGlnThrS 558
2385 GGCACCCAGCCCTCTGTGTTGGACACGACCATCCAGCACCAGCGGCCAG 2434
558 erArThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnPhe 574
2435 CCGCGACCCCTCACACTGATTGCCAAGGTCAATCCAGAACCTCGCACACCGT 2484
575 SerLysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLe 591
2485 GCCCGCTTCGGTGAGAAGGAGCGCTACATGGCTTCATGAATAGCTTCCT 2534
591 uGluLeuGluTrpGlySerMetGlnGlnPheLeuTyTrpGluIleSerAsnL 608
2535 GAGGAACATGAGCACCAGCCATGCAATGCTCTCGACACAGGTAGCCATGG 2584
608 euAspThrLeuThrAsnSerSerPheGluGlyTyTrIleAspLeuGly 624
2585 TGGATGTGGATGCTCCCCCAGTGGTTACCAAGGCACTGGTGTATCTGGCC 2634
625 ArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuProGlnLe 641
2635 CTCACGTTAGCTGCTCATGCCAGCTCTGTGTACAATTTTGTGTGCT 2684
641 uSerLysGluAlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSerA 658
2685 TGACACAGACACCCGAGACCCCTGGAAACCACTGCCACCATCTCTGGAG 2734
658 spIleSerThrAlaLeuArgAsnProAsnIleGlnArgGlnProSerArg 674
2735 CCATTGAG.....GAGGCGCCAGCCGTGTG... 2757
675 GlnSerGluArgAlaArgSerGlnProMetValLeuArgGlyProSerAl 691
2758CTGTGTCGATGCCAATGGCTCTCCCACTGCCCGCCG 2795
691 a.....GlnMetGlnGlyTrpMetMet... 698
2796 CAGGTCACCTCCAGCCCTCCCGCAGGGGAGACCGCGCTCTCTGGCC 2845
699 ..ArgAspLeuAsnSerSerIleAspLeu.....GlnSerPheMetAla 712

[illegible]

OM of: US-09-294-298-4 to: EST.* out_format : pfs

Date: Jan 18, 2001 9:24 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODE=frame+pn2n.model -DEV=xlp
-Q/cgn2/_USPTO.spool/US09294298/runat_17012001_125501_27745/app_query.fasta_1.4359
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-EGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEX=7.000 -START=1 -MATRIX=blossum62 -TRANS=human4.0 cd
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09294298 -CGNJ_1_1719 -NCPU=6
-ICPU=3 -LOGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-294-298-4
Query length: 1376
Database: EST.*
Database sequences: 7991742
Database length: -791223438
Search time (sec): 4688.750000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_est37:AV590698	+ 933.00	1064.14	2.9e-50	571	AV590698 AV590698 Bos taurus b2
gb_gss21:AZ399131	+ 764.00	872.80	1.3e-39	474	AZ399131 IM0164P14R Mouse 10XB
gb_est49:AW79747	- 737.50	839.91	9.0e-38	593	AW79747 hn85f10.x1 NCI_CGAP_K1
gb_est44:AW421227	- 701.00	832.96	2.0e-37	597	AW421227 fj94f10.x1 zebrafish g
gb_est71:BB490923	- 736.00	803.20	1.0e-35	628	BE490923 db38g08.x1 Blackshear
gb_est22:AA106711	+ 660.00	751.72	7.4e-33	566	AI606711 ml58g01.y1 Stratagene
gb_est71:BB463433	- 628.50	717.95	5.6e-31	463	BE463433 hw23g06.x1 NCI_CGAP_K1
gb_est22:AA157567	- 600.00	684.20	4.2e-29	511	AI157567 UT-R-YO-VK-d-04-0-UI-s
gb_est41:AW205989	- 571.50	652.43	2.5e-27	474	AW205989 UI-H-B11-afy-d-12-0-UI
gb_est28:AA398547	+ 566.50	642.16	9.3e-27	707	AJ398547 dkfz426 Gallu
gb_est11:AA709010	- 563.00	643.89	7.5e-27	427	AI609010 zif94hil.s1 Soares_pine
gb_est22:AA1609604	- 548.50	626.44	7.0e-26	460	AI609604 tw91g07.s1 NCI_CGAP_K1
gb_est9:AA554056	- 543.50	621.00	4.4e-25	449	AA554056 n101g07.s1 NCI_CGAP_K1
gb_est27:AA1958829	+ 537.50	611.71	1.6e-25	555	AI958829 fd22g01.y1 zebrafish w
gb_est40:AW136165	- 531.50	607.41	8.0e-25	443	AW136165 UI-H-B11-abz-d-09-0-UI
gb_est10:AA554055	- 517.50	596.16	6.5e-24	452	AA554055 n101g06.s1 NCI_CGAP_K1
gb_est10:AA691827	+ 504.00	576.73	4.1e-23	413	AA691827 vt05g01.r1 Barstead m
gb_gss26:CN502VGM	+ 501.00	563.00	2.4e-22	1020	AL215753 Tetradodon nigroviridis
gb_est74:BB695645	+ 486.00	553.59	8.0e-22	516	BE695645 MRI-BR0800-030700-002
gb_est9:AA575948	- 484.50	555.39	6.3e-22	379	AA575948 nm56g05.s1 NCI_CGAP_K1
gb_est22:AA1562171	- 484.50	555.30	6.3e-22	382	AI562171 vt73f04.x1 Stratagene
gb_est41:AW204612	- 481.00	552.38	9.3e-22	347	AW204612 UI-H-B11-aej-c-02-0-UI
gb_est2:AA107246	+ 480.00	548.54	1.5e-21	440	AA107246 ml58g01.r1 Stratagene
gb_est39:AW076911	+ 473.00	543.29	3.0e-21	345	AW076911 fw39f05.y1 Soares.thym
gb_est76:BB484541	+ 464.00	526.76	2.5e-20	596	BE484541 ul27f03.x1 Sugano.mous
gb_gss26:CN502JDV	+ 461.50	518.25	7.4e-20	979	AL200092 Tetradodon nigroviridis
gb_gss26:CN502JNG	+ 457.50	515.52	1.1e-19	832	AL200329 Tetradodon nigroviridis
gb_gss14:AA0993964	+ 454.00	513.78	1.3e-19	682	AQ993964 RPECI-23-24C10-TV RPEC
gb_gss26:CN501XQ5	+ 452.50	512.95	1.5e-19	631	AI172022 Tetradodon nigroviridis
gb_gss25:CN500HWF	- 425.50	476.73	1.5e-17	1006	AL073826 Drosophila melanogast
gb_gss26:CN502JAH	+ 419.50	470.22	3.5e-17	975	AL235430 Tetradodon nigroviridis
gb_est45:AA527364	- 418.50	481.87	7.9e-18	317	AW527364 UI-R-B01-ajj-e-01-0-UI
gb_est59:BB321064	+ 417.50	479.18	1.1e-17	363	BB321064 BB321064 RIKEN full-le
gb_est26:AA1874961	+ 416.50	478.92	1.2e-17	336	AI874961 ul27f03.x1 Sugano.mous
gb_gss27:CN504DL0	+ 416.50	467.99	4.7e-17	877	AL285885 Tetradodon nigroviridis
gb_est4:AA258366	- 401.50	462.34	9.7e-17	319	AA258366 zf59g01.s1 Soares.Nhm
gb_gss27:CN5041W	+ 401.00	449.63	4.9e-16	926	AI270761 Tetradodon nigroviridis
gb_est47:AA656799	+ 394.00	447.71	6.3e-16	542	AA656799 109245 MARC IBOV Bos t
gb_est61:BB394547	+ 383.50	441.52	1.4e-15	325	BB394547 BB394547 RIKEN full-le
gb_est23:AA1650331	+ 379.00	437.43	2.4e-15	296	AI650331 wal8f01.x1 NCI_CGAP_K1
gb_est14:AA953060	- 368.50	425.57	1.1e-14	292	AA953060 o001g09.s1 Soares_NFL

gb_est59:BB312776 + 365.50 421.49 1.8e-14 309 BB312776 h0312776 1140
gb_est75:BE773015 - 365.50 416.45 3.5e-14 481 BE773015 RCI-FP0144-176
gb_est19:AI327335 - 358.00 404.42 1.6e-13 651 AI327335 mp74c11.x1 Soua
gb_est66:BE085114 - 357.50 411.41 6.6e-14 335 BE085114 CM2-P1.14 2402hu
seq_name: gb_est37:AV590698

seq_documentation_block:

LOCUS AV590698 571 bp mRNA EST 29-AUG-2000
DEFINITION AV590698 Bos taurus brain fetus Bos taurus cDNA clone EIBR0140.1
5', mRNA sequence.
ACCESSION AV590698
VERSION AV590698.1 GI:9701691
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 571)
Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
Suzuki,H.
TITLE bovine cDNA sequencing
JOURNAL Unpublished (2000)
COMMENT Contact: Yoshikazu Sugimoto
Animal Genetics Division
Shirakawa Institute of Animal Genetics
Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8001, Japan
Tel: 81-248-25-5641
Fax: 81-248-25-5725
Email: kazusugi@cocoa.ocn.ne.jp
Single pass sequencing.
This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source
1..571
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="EIBR014B04"
/clone.lib="Bos taurus brain fetus"
/tissue_type="brain"
/dev_stage="fetus"
/lab_host="DH108"
BASE COUNT 124 a 161 c 169 g 116 t 1 others
ORIGIN

alignment_scores:

Quality: 933.00 Length: 186
Ratio: 5.043 Gaps: 1
Percent Similarity: 99.462 Percent Identity: 99.462

alignment_block:

US-09-294-298-4 x AV590698

Align seg 1/1 to: AV590698 from: 1 to: 571

382 LysGluGluValAlaSerAlaLeuValHisIleLeuGlnSerThrGlyLy 398

13 AAGAGGAGGTCGCCAGTCTCTGGTTCACATCCCTGCAGAGTACAGCAA 62

398 sAlaLysAspPheLeuSerAspMetAlaMetSerGluValAspArgPheM 415

63 GGCCAGGAGCTCTTTCAGACATGCCCATGTCCGAGGTGCACCTTCA 112

415 etGluArgGluHisLeuIlePheArgGluAsnThrLeuAlaThr 441

113 TGAACGAGACACCTCATATTCCCGAGAACACCTCCGCCATAAA 162

432 ileGluGlyMetArgLeuIleGlyIleLysTyrLeuLysAspAla 448

163 ATAGAGAGTACATGACATGATTGGTCAGAAATACCTCAAGGATGCCAT 212

448 edlyGluPheIleArgAlaLeuTyrGluSerGluAsnCysGluValA 465
 213 CGGGAAATTCATCCGTCCTCTGTATGAATCTGAGGAACTGTGAGGTGG 262
 465 spProIleLysCysThrAlaSerSerLeuAlaGluHisGlnAlaAsnLeu 481
 263 ACCCCATCAAGTCAGCGCATCCAGTTGGCTGAGCAGCAGGCCAACCTG 312
 482 ArgMetCysCysGluLeuAlaLeuCysLysValValAsnSerHisCysVa 498
 313 CGAGTGTCTGTGAGTTGGCCCTGTCAAGGTGCTCAACTCCCATTCGCT 362
 498 lPheProArgGluLeuLysGluValPheAlaSerTrpArgLeuArgCysA 515
 363 GTTCCGAGGAGGAGCTGAAGGAGGTGTTGGCTCTCGCGCACTCGCTGTG 412
 515 laGluArgGlyArgGluAspIleAlaAspArgLeuIleSerAlaSerLeu 531
 413 CAGACGGGGCGGAGAGACATTCGACCGCTGATCAGCGCTCGCTC 462
 532 PheLeuArgPheLeuCysProAlaIleMetSerProSerLeuPhe.GlyL 548
 463 TTCCCTGCGCTCTCTGCGCGGCATATGTCGCCAGCCCTNTTCGGGGC 512
 548 euMetGluGluTyrProAspGluGlnThrSerArgThrLeuThrLeuIle 564
 513 TCATCAGAGAGTACCCAGATGAGCAAACTTCGAGAACCTTCACGCTCATC 562
 565 AlaLys 566
 563 GCCAAG 568

seq_name: gb_gss21:AZ399131

seq_documentation_block: 474 bp DNA GSS 03-OCT-2000
 LOCUS AZ399131
 DEFINITION IM0164P14 Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 clone UUGCLM0164P14 R, DNA sequence.

ACCESSION AZ399131

VERSION AZ399131.1 GI:10514203

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 474)
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert length: 10000 Std Error: 0.00

Plate: 0164 row: P column: 14

Seq primer: CACACAGGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 474.

Location/Qualifiers

FEATURES

SOURCE

1..474
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCLM0164P14"

seq_name: gb_est49:AW779747

/clone_lib="Mouse 10kb plasmid UUGCLM library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, P-"
 /note="vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of PWD42 (gii4732114[gb]AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."
 BASE COUNT 92 a 151 c 131 g 100 t
 ORIGIN

alignment_scores:
 Quality: 764.00 Length: 142
 Ratio: 5.380 Gaps: 0

Percent similarity: 100.000 Percent identity: 99.296

alignment_block:

US-09-294-298-4 x AZ399131/rev ..

Align seg 1/1 to reverse of: AZ399131 from: 1 to: 474

192 LysAspAsnSerArgArgValAspAsnValLeuLysLeuTrpIleLeG 208

:::|||||

426 CAGGCAACAGCGCGCGAGTAGATACGCTGGAAGCTATGGATCATAGA 377

208 uAlaArgGluLeuProProLysLysArgTyrTyrCysGluLeuCysLeuA 225

|||||

376 GGCTCGAGAGCTGCCCGCCCAAGAGAGATATTACTGTGAGCTGTGCCTGG 327

225 spAspMetLeuTyrAlaArgThrSerLysProArgSerAlaSerGly 241

|||||

326 ACACATGCTGTATGCACCAACCACTCCAGCCCGCTCGGCTTCAGGA 277

242 AspThrValPheTrpGlyGluHisPheGluPheAsnAsnLeuProAlaVa 258

|||||

276 GACACCGCTCTTTGGGGCGAGCACCTTTGAGTTTAAACACCTGCCTGCCGT 227

258 lArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysLysL 275

|||||

226 CCGGGCCCTTCGCGTCATCTCTACCGTCACTCAGACAAAAAGCGGAAGA 177

275 ysAspLysAlaGlyTyrValGlyLeuValThrValProValAlaThrLeu 291

|||||

176 AGGACAAGCGTGGCTAGCTTGGCTGGTGGCTGCTTCCAGTGGCCACCTA 127

292 AlaGlyArgHisPheThrGluGlnTrpTyrProValThrLeuProThrG 308

|||||

126 GCTGGCGGCCACTTCACAGAGCAGTGGTACCCCGTGACCTTGGCGACAGG 77

308 ySerGlyGlySerGlyGlyMetGlySerGlyGlyGlyGlyGlyGlyG 325

|||||

76 CAGTGGGGGCTCTGGGGGCATGGCTCGGGGGAGGAGGAGGGTTCAGGGG 27

325 lyGlySerGlyGlyLysGlyLysGly 333

|||||

26 GTGGCTCAGGGGGCAAGGGAAGGA 1

score_list:	Sequence	Strid Orig	zScore	EScore Len	Documentation	...
	/SID36/cgcgdata/geneseq/NA1992.DAT:Q26078 +			437.50	304.80	3.4e-09
	/SID36/cgcgdata/geneseq/NA1991.DAT:Q10988 +			435.50	303.40	4.1e-09
	/SID36/cgcgdata/geneseq/NA1991.DAT:T76739 +			429.50	302.84	4.4e-09
	/SID36/cgcgdata/geneseq/NA1998.DAT:T62176 +			311.50	192.36	0.0063
	/SID36/cgcgdata/geneseq/NA2000.DAT:T82718 +			298.50	198.15	0.0030
	/SID36/cgcgdata/geneseq/NA1997.DAT:T68715 +			293.50	194.67	0.0047
	/SID36/cgcgdata/geneseq/NA2000.DAT:T58980 +			286.00	194.56	0.0048
	/SID36/cgcgdata/geneseq/NA1999.DAT:T23430 +			283.00	197.09	0.0034
	/SID36/cgcgdata/geneseq/NA1998.DAT:T662154 -			274.50	179.32	0.0036
	/SID36/cgcgdata/geneseq/NA1995.DAT:T86129 +			273.00	189.03	0.0097
	/SID36/cgcgdata/geneseq/NA1995.DAT:T57917 +			273.00	187.42	0.0119
	/SID36/cgcgdata/geneseq/NA2000.DAT:T87351 +			270.00	172.03	0.0856
	/SID36/cgcgdata/geneseq/NA1996.DAT:T06769 +			263.50	169.37	0.1205
	/SID36/cgcgdata/geneseq/NA1997.DAT:T89595 +			263.50	169.27	0.1219
	/SID36/cgcgdata/geneseq/NA1997.DAT:T05287 +			263.50	165.31	0.2028
	/SID36/cgcgdata/geneseq/NA1999.DAT:T584636 +			260.50	180.86	0.0276
	/SID36/cgcgdata/geneseq/NA1997.DAT:T80413 -			259.50	163.50	0.2258
	/SID36/cgcgdata/geneseq/NA1997.DAT:T80518 +			253.50	159.12	0.4482
	/SID36/cgcgdata/geneseq/NA1997.DAT:T80414 -			253.50	159.12	0.4482
	/SID36/cgcgdata/geneseq/NA1992.DAT:T29263 +			253.00	174.38	0.0633
	/SID36/cgcgdata/geneseq/NA1999.DAT:T06978 +			252.00	176.68	0.0471
	/SID36/cgcgdata/geneseq/NA1999.DAT:T21373 +			252.00	176.68	0.0471
	/SID36/cgcgdata/geneseq/NA1999.DAT:T81748 +			251.50	174.82	0.0599
	/SID36/cgcgdata/geneseq/NA1995.DAT:T808016 +			251.00	185.47	0.0153
	/SID36/cgcgdata/geneseq/NA1999.DAT:T36583 +			251.00	170.19	0.1084
	/SID36/cgcgdata/geneseq/NA1992.DAT:T02602 +			251.00	169.29	0.1216
	/SID36/cgcgdata/geneseq/NA1997.DAT:T46941 +			251.00	167.95	0.1445
	/SID36/cgcgdata/geneseq/NA1997.DAT:T80415 -			251.00	165.96	0.1864
	/SID36/cgcgdata/geneseq/NA1999.DAT:T584643 +			250.50	173.48	0.0711
	/SID36/cgcgdata/geneseq/NA1997.DAT:T86704 -			250.00	176.36	0.0492
	/SID36/cgcgdata/geneseq/NA1995.DAT:T96261 +			249.00	175.49	0.0550
	/SID36/cgcgdata/geneseq/NA1994.DAT:T64201 +			248.50	171.31	0.0939
	/SID36/cgcgdata/geneseq/NA1997.DAT:T89345 +			247.50	173.02	0.0755
	/SID36/cgcgdata/geneseq/NA1995.DAT:T84659 +			247.50	167.87	0.1452
	/SID36/cgcgdata/geneseq/NA1995.DAT:T084659 +			247.50	167.86	0.1452
	/SID36/cgcgdata/geneseq/NA1994.DAT:T73500 -			247.50	167.28	0.1575
	/SID36/cgcgdata/geneseq/NA2000.DAT:T253887 -			246.50	151.69	1.16
	/SID36/cgcgdata/geneseq/NA1991.DAT:T13669 +			246.50	174.38	0.0633
	/SID36/cgcgdata/geneseq/NA1999.DAT:T62134 -			246.50	158.12	0.5097
	/SID36/cgcgdata/geneseq/NA1992.DAT:T24828 +			245.50	165.74	0.1918

147 valHisSerI

Align seg 1/1 to: Q26078 from: 1 to: 4307

147 valHisSerSerIleLeuGlyGlnGluPheCysPheGluValThrSer 163

443 rLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTyrGlnSerGluG 460
2545 TGTTCATCATGCTTTGAAGAGCTCTATTTTAAAGAGTAATGGAACGAAGC 2594
460 LuAsnCysGluValAspProIleLysCysThrAlaSerSerLeuAlaGlu 476
2595 AGTCYGTGGAGTAASTCCATAAGTTAGAAAAAATGAAGATGTCGAC 2644
477 HisGlnAlaAsn.....LeuArgMetCysCysGluLeuAlaLeuCysLy 491
2645 ACTAATTTAAACACACACTTATGAACATACCTTCACAGAGCTT...GTGGAGAA 2691
491 sValValAsnSerHisCysValPheProArgGluLeuLysGluValPheA 508
2692 AATATTATGCGCTTCAGAAATACITCCACCACCATTCAGATATATTTATG 2741
508 laSerTrpArgLeuArgCysAlaGluArgGly..... 518
2742 GG.....TGTTCACAGAATCTGTTACGACATAAGTGGCCT 2776
519 ...ArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuAr 534
2777 ACAAAATACCACCATGAGAACCAAGACTGTTAGTGGTGTGTTTTCTTCG 2826
534 oPheLeuCysProAlaIleMetSerProSerPheLeuPheGlyLeuMetGlnG 551
2827 ACTCATCTGCTCGCATCTCGAATCCAGCATGTTCAATATCATCTACG 2876
551 LuTyrProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysVal 567
2877 ATTCTCCATCTCCTATTGCTCCAAAGACACTGATATTAGTGGCTAAATCT 2926
568 IleGlnAsnLeuAlaAsnPheSerLysPheThrSerLysGluAspPheLe 584
2927 GTGCAGAACTTAGCAATCTTGCGAATTTGGAGCTAAGGAGCGCCTACAT 2976
584 uGlyPheMetAsnGluPheLeuGluLeuGluTyrPGlySerMetGlnGlnP 601
2977 GGAAGGTGTCAATCCATTATCAAAAGCAACAAACATCGTATGATCATGT 3026
601 heLeuTyrGlnIleSerAsnLeuAspThrLeuThrAsnSerSerSerPhe 617
3027 TTTTAGATGAACCTTGGGAATGTACCTGAACCTTCGGACACTACAGAG... 3073
618 GluGlyTyrIleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLe 634
3074 CATCTTAGACCGGACTGTCCCGTGATTTACGCATTTGCATGAGATTG 3123
634 u.....TrpGluValLeuProGlnLeuSerLysGlu..... 644
3124 CGTGGCTCATTTACAGATGAACITTCGAACGCTCAGTAATGAGCGTGGTGAC 3173
645AlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSerAsp 658
3174 AGCAGCAGCTATTGAAAAGCTTCTGGCTATATACAGAACTGCTTCAACAA 3223
659 IleSerThrAlaLeuArgAsnProAsn...IleGlnArgGlnProSerAr 674
3224 AAACAAAACCGATATACAAAACCAATGATGTCAGGTAGCAGCCTTCGCC 3273
674 gGln 675
3274 CCAG 3277

DT 22-MAY-1991 (first entry)

OM of: US-09-294-298-4 to: Issued_Patents_NA:* out_format: pfs
 Date: Jan 19, 2001 7:48 AM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+psn.model -DRV=xlsp
 -Q=/cgn2_1/USP90.spool/US09294298/runat_17012001_125501_27768/app_query.fasta_1.4359
 -DB=Issued_Patents_NA -QFMT=fastp -SUFFIX=rni -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -GAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -GAPEXT=6.000 -XGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -MINLEN=0 -MAXLEN=200000000
 -USER=US09294298_CGMA_1_163 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
 -WAIT -THREADS=1

Search information block:

Query: US-09-294-298-4
 Query length: 1376
 Database: Issued_Patents_NA:*
 Database sequences: 280836
 Database length: 80580151
 Search time (sec): 156.930000

score.list:

Sequence	Strd Orig	ZScore	Len	Documentation
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-24 +	437.50	421.08	6.7e-16	3456
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-7 +	437.50	418.91	8.8e-16	4307
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-909-954-3 +	371.00	355.38	3.0e-12	3117
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-909-954-1 +	362.50	346.76	9.2e-12	3148
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-557-139-1 +	283.00	260.40	5.9e-07	6152
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-258-261B-6 +	263.50	225.58	5.2e-05	28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-456-837-6 +	263.50	225.58	5.2e-05	28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-342-6 +	263.50	225.58	5.2e-05	28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-646A-6 +	263.50	225.58	5.2e-05	28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-076A-6 +	263.50	225.58	5.2e-05	28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-764-233A-6 +	263.50	225.58	5.2e-05	28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-339A-6 +	263.50	225.58	5.2e-05	28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-729-214-6 +	263.50	225.58	5.2e-05	28958
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-09-028-934-6 +	263.50	225.58	5.2e-05	28958
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-764-233A-1 +	263.50	220.32	0.0001	49377
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-1 +	259.50	217.61	0.0001	43280
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-7 +	253.50	211.34	0.0003	44377
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-198-1 +	253.00	211.34	0.0003	44377
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-781-802-7 +	253.00	236.91	1.2e-05	3147
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-551-531-1 +	251.00	228.45	3.6e-05	6060
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-449-933-1 +	251.00	224.62	5.8e-05	8937
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-411-389-1 +	251.00	222.84	7.3e-05	10706
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-13 +	251.00	220.21	0.0001	13987
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-574-959A-6 +	248.50	230.28	2.8e-05	3901
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-403-852D-1 +	248.50	227.09	4.3e-05	5392
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-510-646B-1 +	248.50	227.09	4.3e-05	5392
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-949-386-23 +	247.50	222.46	7.7e-05	7791
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-450-562-23 +	247.50	222.46	7.7e-05	7791
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-949-386-22 +	247.50	222.44	7.7e-05	7808
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-450-562-22 +	247.50	222.44	7.7e-05	7808
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-945-283-1 +	247.50	221.67	8.5e-05	8438
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-09-335-409-1 +	247.50	201.01	0.0012	68750
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-149-097D-22 +	247.00	221.94	8.2e-05	7808
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-7 +	245.00	202.81	0.0010	44377
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-198-1 +	245.00	202.81	0.0010	44377
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-779-801-3 +	243.50	235.96	4.9e-05	3632
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-343-785-7 +	243.50	230.59	2.7e-05	2051
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-462-221-7 +	242.50	230.59	2.7e-05	2051
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-946-458-7 +	242.50	230.59	2.7e-05	2051
/cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-343-785-1 +	242.50	230.31	2.8e-05	2111

/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-462-221-1 + 242.50 230.31 2.8e-05 2111
 /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-946-458-1 + 242.50 230.31 2.8e-05 2111
 /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-574-959A-8 + 242.00 225.67 5.1e-05 3211
 /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-14418-1 + 241.50 225.69 5.1e-05 3211
 /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-15327-1 + 241.50 225.69 5.1e-05 3211
 seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-24

seq_documentation_block:
 ; Sequence 24, Application US/08190687B
 ; Patent No. 5760203
 ; GENERAL INFORMATION:
 ; APPLICANT: Wong, Gail L.
 ; APPLICANT: Martin, George
 ; APPLICANT: McCormick, Francis P.
 ; APPLICANT: Rubinfeld, Bonnie
 ; APPLICANT: O'Rourke, Edward C.
 ; APPLICANT: Clark, Robin
 ; TITLE OF INVENTION: GAP Gene Sequences
 ; NUMBER OF SEQUENCES: 27
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/190,687B
 ; FILING DATE: 02-FEB-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/774,644
 ; FILING DATE: 11-OCT-1991
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/260,807
 ; FILING DATE: 21-OCT-1988
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/230,761
 ; FILING DATE: 10-AUG-1988
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Gass, David A.
 ; REGISTRATION NUMBER: 38,153
 ; REFERENCE/DOCKET NUMBER: 27527/31898
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 24:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3456 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 100..2709
 ; US-08-190-687B-24

alignment_scores:
 Quality: 437.50 Length: 568
 Ratio: 1.326 Gaps: 22
 Percent Similarity: 58.099 Percent Identity: 26.232
 alignment_block:
 US-09-294-298-4 x US-08-190-687B-24

Align seg 1/1 to: US-08-190-687B-24 from: 1 to: 3456

```

147 ValHisSerSerIleLeuGlyGlnGluPheCysPheGluValThrThr3e 163
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1156 GTTCATGATAGTCTCTTTGGCAGGCCAACTGTTTCAGATAGTCTCA 1205
163 rSerGlyThrLys.....CysPheAlaCysArgSerAlaA 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1206 GCACCTTTAGTGAAGAACATTACATCTTTTACTTTGACGAGAACTCCAG 1255
175 laGluArgAspLysTrpIleGluAsnLeuGln.....Arg 186
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
1256 AACAGCAGCAGATTGGATGAAGAGCTGTCAGGCAATTTTACG 1305
187 AlaValLysProAsnLysAspAsnSerArg...ArgValAspAsnValLe 202
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1306 AAAAGTAGTCCAGGACATCCCAATAAACGCCCTTCGTGAGTCAGCCT 1355
202 uLysLeuTrpIleIleGluAlaArgGluLeuProProLys.....LysA 217
| ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:
1356 TGTTTTACATATTGAAGAGGCCATAAATCCCAATAAACATTTTACTA 1405
217 rGlyTrpCysGluLeuCysLeuAspAspMetLeuTrpAlaArgThrThr 233
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
1406 ATCCATATTGTAACATCTACCTGAATAGTGTCCAGTAGCAAAACTCAT 1455
234 SerLysProArgSerAlaSerGlyAspThrValPheTrpGlyGluHisph 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1456 GCAAGG.....GAAGGCCAAACCCAGTATGGTCGAAGAGTT 1493
250 eGluPheAsnAsnLeuProAla...ValArgAlaLeuArgLeuHisLeuT 266
| ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:
1494 TGTCTTTGATGATCTTCTCTGACATCAATAGATTGAAATAACTCTT. 1542
266 yArgAspSerAspLysLysArgLysLysAspLysAlaGlyTrpValGly 282
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
1543 .....AGTAAATAAAACAAAGAAAGCAAGATCTGTATCTTA 1581
283 LeuValThrValProValAlaThrLeuAlaGlyArgHisPheThrGluG 299
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1582 TTTATGCGTGGCAGTTCGCGGATTACAGAAAGGCCATGCCACAGATGA 1631
299 nTrpTrpProValThr.....LeuProThrGlySerGlyGlySerGlyG 314
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1632 ATGGTTCTGCTCAGTCCCATATACCATTA..... 1662
314 LyMetGlySerGlyGlyGlyGlySerGlyGlyGlySerGlyGlyLys 330
1662 ..... 1662
331 GlyLysGlyGlyCysPro...AlaValArgLeuLysAlaArgTrpGlnTh 346
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1663 .....AAAGTATTGAACAGGTCCTCGGTGTCGACGACGATCTAT 1709
346 rMetSerIleLeuProMetGluLeuTrpLysGluPheAlaGluTrpValT 363
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
1710 GGAATAAATCATGCCAGAAGAGAGTACAGTGAATTTAAAGAGCTTATAC 1759
363 hrAsnHisTrpArgMetLeuCysAlaValLeuGluProAlaLeuAsnVal 379
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
1760 TCAGAAAGGAATTCATGTAGTCTATGCTTATCA.....CATGTA 1800
380 LysGlyLysGluGlu.....ValAlaSerAlaLeuValHisIleLeuG 394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1801 TGTGACAAAGCCGAACACTACTGCCAGCATCTCTACTGAGGATTTTCT 1850
394 nSerThrGlyLysAlaLysAspPheLeuSerAspMetAlaMetSerGluV 411
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
1851 TCACGAAAAGCTTGAATCGTTGTTGTTATGCACACTAAATGACAGAGAAA 1900
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
411 alaAspArgPheMetGluGlu...HisLeuIlePheArgGluAsnThr 426
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

1901 TAAGC.....ATGGAAGATGAAGCCACTACCTATTTCGAGCCACAACA 1944
427 LeuAlaThrLysAlaIleGluLeuTrpMetArgLeuIleGlyGlnLysTy 443
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1945 CTTGCAAGCACCCTTGATGAGCAGTATATGAAGCCACTGCTACACAGTT 1994
443 rLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTyGluSerGluG 460
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1995 TGTTCATCATGCTTTGAAGAGACTCTATTTTAAAGATATATGGAAGCAAGC 2044
460 luAsnCysGluValAspProIleLysCysThrAlaSerSerLeuAlaGlu 476
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
2045 AGTCTTGTGAGTTAAGTCCATCAAGATTAGAAAAAATAGAGATGTGAAC 2094
477 HisGlnAlaAsn.....LeuArgMetCysGlyGluLeuAlaLeuCysLy 491
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
2095 ACTAATTTAACACACCTATTGACATACCTTTTCAGAGCTT...GTGGAGAA 2141
491 sValValAsnSerHisCysValPheProArgGluLeuLysGluValPheA 508
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2142 AATATTTCATGGCTTCAGAAATACTTCCACCGACATTCAGATATATTATG 2191
508 laSerTrpArgLeuArgCysAlaGluArgGly..... 518
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
2192 GG.....TGTTTACAGAAATCTGTTACGATTAAGTGCCT 2226
519 ...ArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuAr 534
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
2227 ACAATATACCACCATGAGAACAGAGTGTGTAGTGGTTTGTGTTCTTCG 2276
534 gPheLeuCysProAlaIleMetSerProSerLeuPheGlyLeuMetGlnG 551
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2277 ACTCATCTGTCTGCTCCATCTGTAATCCAGGATGTTCAATATCATCTCAG 2326
551 luTrpProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysVal 567
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2327 ATTCTCCATCTCTTATGTCGCAAGACACTGATATTAGTGGCTAAATCT 2376
568 IleGlnAsnLeuAlaAsnPheSerLysPheThrSerLysGluAspPheLe 584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2377 GTCAGAACTTAGCAAACTCTGTGGAATTTGGAGCTAAGGAGCCCTACAT 2426
584 uGlyPheMetAsnGluPheLeuGluLeuTrpGlySerMetGlnGlnP 601
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2427 GGNAGGTGCATTCATTCATCAAAAGCAACAAACATCGTATGATCATGT 2476
601 heLeuTrpGluIleSerAsnLeuAspThrLeuThrAsnSerSerPhe 617
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2477 TTTAGTAGAATTTGGGAATGTACCTGAACCTCCGACACTACAGAG... 2523
618 GluGlyTrpIleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLe 634
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
2524 CATCTAGAACCGACCTGTCCCGTGAATTTAGCAGCATTCGATGAGATTG 2573
634 u.....TrpGluValLeuProGlnLeuSerLysGlu..... 644
2574 CFTGGCTCATTCAGATGAACCTCGAACCGCTCAGTAATGAGCGTGGTGCAC 2623
645 .....AlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSerAsp 658
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
2624 AGCAGCAGCTATTGAAAAGCTTCTGGCTATAACAGAACTGCTTCAACAA 2673
659 IleSerThrAlaLeuArgAsnProAsn...IleGlnArgGlnProSerAr 674
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2674 AACAAAAACCACTATACAAAAACCAATGATGTAGTAGCAGCGCTTCGCC 2723
674 gGln 675
|||
2724 CCAG 2727

```

seq_name: /cgn2_6/ptodata/2/ina/5A_comb.seq:US-08-190-687B-7

seq_documentation_block: